SOME COMPUTATIONAL AND MODEL EQUIVALENCES IN ANALYSES OF VARIANCE OF UNEQUAL-SUBCLASS-NUMBERS DATA

BU-668-M by March, 1979

S. R. Searle, F. M. Speed[1/] and H. V. Henderson
Biometrics Unit, Cornell University, Ithaca, New York

Abstract

Available methodologies for calculating sums of squares in analyses of variance of unequal-subclass-numbers (unbalanced) data include (i) full rank reparameterized models, (ii) "indirect" methods, (iii) the $R(\cdot|\cdot)$ notation, (iv) weighted squares of means and (v) numerator sums of squares for testing hypotheses. These techniques are described, and relationships between them explained and illustrated. Numerical illustrations are given in the Appendix.

1. Introduction

A variety of methodologies for calculating sums of squares in the analyses of variance of unequal-subclass-numbers (unbalanced) data are currently available in the literature and are being used in computer programs for such analyses. Originating with balanced data analyses, there is the use of restrictions such as $\Sigma\alpha_i = 0$ and $\Sigma\beta_j = 0$ to reparameterize over-parameterized models to make them of full rank. There is also a methodology that involves inverting submatrices of the inverse of an $\mathbf{X}'\mathbf{X}$ matrix in full rank models $E(\mathbf{y}) = \mathbf{Xb}$, a methodology described

[1/] Department of Computer Science and Statistics, Mississippi State University, Mississippi State, Mississippi.

Paper No. BU-668-M in the Biometrics Unit Mimeo Series, Cornell University.

in Henderson [1959, 1969] and Searle [1966], and referred to in Harvey [1970] as the "indirect method". There are also the sums of squares from the weighted squares of means analyses, first suggested by Yates [1934] and more recently available in Gosslee and Lucas [1965] and Searle [1971, p. 369]. There is the $R(\cdot|\cdot)$ notation emphasized in Searle [1971, 1972], and extended by Speed et al. [1978] and Hocking et al. [1978] to cover reparameterized models. And there is the direct hypothesis-testing approach of calculating the numerator sum of squares for testing a linear hypothesis (e.g., Searle [1971, p. 190]). With at least these five different ways of looking at essentially the same calculation, we feel the time is ripe for showing some of the relationships between them; and this is the object of this paper.

We deal throughout with familiar linear models of the sort $E(\underset{\sim}{y}) = \underset{\sim}{X}\underset{\sim}{b}$ where $\underset{\sim}{y}$ is a vector of observations with expected value $\underset{\sim}{X}\underset{\sim}{b}$ over repeated sampling, where $\underset{\sim}{b}$ is a vector of unknown parameters and $\underset{\sim}{X}$ is a known matrix. Least squares estimation of $\underset{\sim}{b}$ leads to normal equations

$$\underset{\sim}{X}'\underset{\sim}{X}\hat{\underset{\sim}{b}} = \underset{\sim}{X}'\underset{\sim}{y} \; . \tag{1}$$

Unbiased estimation of the variance of the assumed homoscedastic y's is taken as $\hat{\sigma}^2 = [\underset{\sim}{y}'\underset{\sim}{y} - R(\underset{\sim}{b})]/[N - r(\underset{\sim}{X})]$, where $R(\underset{\sim}{b})$ is the reduction in sum of squares due to fitting the model, where N is the number of observations (the order of the vector $\underset{\sim}{y}$) and $r(\underset{\sim}{X})$ is the rank of $\underset{\sim}{X}$. Much of the paper is concerned with calculating $R(\underset{\sim}{b})$ in a variety of forms and situations. General development is given in the main body of the paper, with references therein to numerical illustrations given in the Appendix.

2. Full Rank Models

2.1. General results

Models $E(\underset{\sim}{y}) = \underset{\sim}{X}\underset{\sim}{b}$ in which $\underset{\sim}{X}$ has full column rank have normal equations to which the solution is $\hat{\underset{\sim}{b}} = (\underset{\sim}{X}'\underset{\sim}{X})^{-1}\underset{\sim}{X}'\underset{\sim}{y}$, and the reduction in sum of squares for fitting such a model is

$$R(\underset{\sim}{b}) = \underset{\sim}{y}'\underset{\sim}{X}(\underset{\sim}{X}'\underset{\sim}{X})^{-1}\underset{\sim}{X}'\underset{\sim}{y} = \hat{\underset{\sim}{b}}'\underset{\sim}{X}'\underset{\sim}{y} . \tag{2}$$

We refer to $\hat{\underset{\sim}{b}}'\underset{\sim}{X}'\underset{\sim}{y}$ as the R-algorithm for calculating $R(\underset{\sim}{b})$; it consists of the inner product of the vector of solutions with the vector of right-hand-sides of the normal equations (1). We utilize the R-algorithm in the sequel.

2.2. Partitioning the model

Partitioning of $\underset{\sim}{b}'$ as $\underset{\sim}{b}' = [\underset{\sim}{b}_1' \quad \underset{\sim}{b}_2']$ is often appropriate in which case, with the conformable partitioning of $\underset{\sim}{X}$ as $\underset{\sim}{X} = [\underset{\sim}{X}_1 \quad \underset{\sim}{X}_2]$, we get the reduction in sum of squares as

$$R(\underset{\sim}{b}_1, \underset{\sim}{b}_2) = \underset{\sim}{y}'[\underset{\sim}{X}_1 \quad \underset{\sim}{X}_2] \begin{bmatrix} \underset{\sim}{X}_1'\underset{\sim}{X}_1 & \underset{\sim}{X}_1'\underset{\sim}{X}_2 \\ \underset{\sim}{X}_2'\underset{\sim}{X}_1 & \underset{\sim}{X}_2'\underset{\sim}{X}_2 \end{bmatrix}^{-1} \begin{bmatrix} \underset{\sim}{X}_1' \\ \underset{\sim}{X}_2' \end{bmatrix} \underset{\sim}{y} . \tag{3}$$

Similarly for fitting just part of the model, namely $E(\underset{\sim}{y}) = \underset{\sim}{X}_2\underset{\sim}{b}_2$, the reduction in sum of squares is

$$R(\underset{\sim}{b}_2) = \underset{\sim}{y}'\underset{\sim}{X}_2(\underset{\sim}{X}_2'\underset{\sim}{X}_2)^{-1}\underset{\sim}{X}_2'\underset{\sim}{y} ,$$

with the difference between these two being denoted by $R(\underset{\sim}{b}_1|\underset{\sim}{b}_2)$:

$$R(\underset{\sim}{b}_1|\underset{\sim}{b}_2) = R(\underset{\sim}{b}_1, \underset{\sim}{b}_2) - R(\underset{\sim}{b}_2) . \tag{4}$$

This represents the sum of squares that is often called the sum of squares due to $\underset{\sim}{b}_1$ adjusted for $\underset{\sim}{b}_2$.

On adopting the usual normality assumptions, it is well recognized that $R(\underset{\sim}{b}_1|\underset{\sim}{b}_2)$ defined by (4) is identical to the numerator sum of squares for testing the hypothesis $H: \underset{\sim}{b}_1 = \underset{\sim}{0}$, using an F-statistic in which the denominator is $r_1\hat{\sigma}^2$, where r_1 is the order of $\underset{\sim}{b}_1$. We will use the phrase

$$\text{using } R(b_1|b_2) \text{ tests } H : \underset{\sim}{b}_1 = \underset{\sim}{0} \tag{5}$$

where, throughout this paper, by the word "using" in this context we mean "using as a numerator sum of squares" in the aforementioned manner.

Suppose $\hat{\underset{\sim}{b}}_1$ and $\hat{\underset{\sim}{b}}_2$ are solutions to the normal equations for $\underset{\sim}{b}_1$ and $\underset{\sim}{b}_2$:

$$\begin{bmatrix} \hat{\underset{\sim}{b}}_1 \\ \hat{\underset{\sim}{b}}_2 \end{bmatrix} = (\underset{\sim}{X}'\underset{\sim}{X})^{-1}\underset{\sim}{X}'\underset{\sim}{y} = \begin{bmatrix} \underset{\sim}{X}_1'\underset{\sim}{X}_1 & \underset{\sim}{X}_1'\underset{\sim}{X}_2 \\ \underset{\sim}{X}_2'\underset{\sim}{X}_1 & \underset{\sim}{X}_2'\underset{\sim}{X}_2 \end{bmatrix}^{-1} \begin{bmatrix} \underset{\sim}{X}_1'\underset{\sim}{y} \\ \underset{\sim}{X}_2'\underset{\sim}{y} \end{bmatrix} . \tag{6}$$

Then it can be shown (e.g., Searle [1971, p. 115]) that

$$\text{using } Q_1 = \hat{\underset{\sim}{b}}_1'\underset{\sim}{T}_{11}^{-1}\hat{\underset{\sim}{b}}_1 \text{ tests } H : \underset{\sim}{b}_1 = \underset{\sim}{0} \; . \tag{7}$$

where $\underset{\sim}{T}_{11}$ is defined as that part of $(\underset{\sim}{X}'\underset{\sim}{X})^{-1}$ which corresponds to $\underset{\sim}{X}_1'\underset{\sim}{X}_1$; i.e.,

$$\begin{bmatrix} \underset{\sim}{T}_{11} & \underset{\sim}{T}_{12} \\ \underset{\sim}{T}_{21} & \underset{\sim}{T}_{22} \end{bmatrix} \equiv \begin{bmatrix} \underset{\sim}{X}_1'\underset{\sim}{X}_1 & \underset{\sim}{X}_1'\underset{\sim}{X}_2 \\ \underset{\sim}{X}_2'\underset{\sim}{X}_1 & \underset{\sim}{X}_2'\underset{\sim}{X}_2 \end{bmatrix}^{-1} . \tag{8}$$

Hence, in this <u>full</u> <u>rank</u> model with $\underset{\sim}{T}_{11}^{-1} = \underset{\sim}{X}_1'\underset{\sim}{X}_1 - \underset{\sim}{X}_1'\underset{\sim}{X}_2(\underset{\sim}{X}_2'\underset{\sim}{X}_2)^{-1}\underset{\sim}{X}_2'\underset{\sim}{X}_1$,

$$Q_1 \equiv \hat{\underset{\sim}{b}}_1'\underset{\sim}{T}_{11}^{-1}\hat{\underset{\sim}{b}}_1 = R(\underset{\sim}{b}_1|\underset{\sim}{b}_2) \; . \tag{9}$$

The form $\hat{\underset{\sim}{b}}_1'\underset{\sim}{T}_{11}^{-1}\hat{\underset{\sim}{b}}_1$ is used by Henderson [1959, 1969]; it is called the "invert part of the inverse" rule and its derivation is shown in Searle [1966, Sec. 9.11, and 1971, Sec. 3.6c]. It is also called the "indirect method" of calculating a numerator sum of squares (for an F-statistic) by Harvey [1970].

3. <u>Non-Full Rank Models</u>

<u>3.1. General results</u>

Models $E(\underset{\sim}{y}) = \underset{\sim}{X}\underset{\sim}{b}$ in which $\underset{\sim}{X}$ has less than full column rank utilize a generalized inverse of $\underset{\sim}{X}'\underset{\sim}{X}$ for solving the normal equations as

$$\underset{\sim}{b}^{o} = (\underset{\sim}{X}'\underset{\sim}{X})^{-}\underset{\sim}{X}'\underset{\sim}{y}$$

where $(\underset{\sim}{X}'\underset{\sim}{X})^{-}$ is such that $\underset{\sim}{X}'\underset{\sim}{X}(\underset{\sim}{X}'\underset{\sim}{X})^{-}\underset{\sim}{X}'\underset{\sim}{X} = \underset{\sim}{X}'\underset{\sim}{X}$. Then, analogous to (2), the reduction in sum of squares is

$$R(\underset{\sim}{b}) = \underset{\sim}{y}'\underset{\sim}{X}(\underset{\sim}{X}'\underset{\sim}{X})^{-}\underset{\sim}{X}'\underset{\sim}{y} = \underset{\sim}{b}^{o'}\underset{\sim}{X}'\underset{\sim}{y} \ . \tag{10}$$

The middle expression in (10) has the same form as that in (2), indeed (2) is the special case of (10) when $\underset{\sim}{X}'\underset{\sim}{X}$ is non-singular; and the $\underset{\sim}{b}^{o'}\underset{\sim}{X}'\underset{\sim}{y}$ of (10) is precisely the R-algorithm for this, the non-full rank model, as is (2) for the full rank model. Appendix equations (A2)-(A4) afford numerical illustration of these points.

Although not every generalized inverse of $\underset{\sim}{X}'\underset{\sim}{X}$ is symmetric it is always possible to construct one that is, in the form $(\underset{\sim}{X}'\underset{\sim}{X})^{-} = \underset{\sim}{G}\underset{\sim}{X}'\underset{\sim}{X}\underset{\sim}{G}'$ where $\underset{\sim}{G}$ is <u>any</u> generalized inverse. $(\underset{\sim}{X}'\underset{\sim}{X})^{-}$ of this form also satisfies $(\underset{\sim}{X}'\underset{\sim}{X})^{-}\underset{\sim}{X}'\underset{\sim}{X}(\underset{\sim}{X}'\underset{\sim}{X})^{-} = (\underset{\sim}{X}'\underset{\sim}{X})^{-}$, and we call it a symmetric reflexive generalized inverse. [(A4) is a symmetric $(\underset{\sim}{X}'\underset{\sim}{X})^{-}$.]

3.2. Hypothesis testing

In non-full rank models, if $\underset{\sim}{K}'$ is a matrix of full row rank s with $s \le r(\underset{\sim}{X})$, and if $\underset{\sim}{K}' = \underset{\sim}{T}'\underset{\sim}{X}$ for some matrix $\underset{\sim}{T}'$ (i.e., if all elements of $\underset{\sim}{K}'\underset{\sim}{b}$ are estimable), then the hypothesis

$$H: \underset{\sim}{K}'\underset{\sim}{b} = \underset{\sim}{m}$$

is tested by (e.g., Searle [1971, p. 192])

$$F(H) = Q/s\hat{\sigma}^2 \quad \text{where} \quad Q = (\underset{\sim}{K}'\underset{\sim}{b}^{o} - \underset{\sim}{m})'(\underset{\sim}{K}'\underset{\sim}{G}\underset{\sim}{K})^{-1}(\underset{\sim}{K}'\underset{\sim}{b}^{o} - \underset{\sim}{m}) \ .$$

Q is the numerator sum of squares of the F-statistic F(H), and in the context described earlier we say

$$\text{using Q tests } H: \underset{\sim}{K}'\underset{\sim}{b} = \underset{\sim}{m} \ . \tag{11}$$

In the special (but oft-used) case of $\mathbf{m} = \mathbf{0}$

$$Q = \mathbf{b}^{o\prime}\mathbf{K}(\mathbf{K}'\mathbf{G}\mathbf{K})^{-1}\mathbf{K}'\mathbf{b}^{o} . \tag{12}$$

These expressions are, of course, equally available to full rank models as they are to those not of full rank.

3.3. Partitioning the model

Partitioning $\mathbf{b}$ as previously, $\mathbf{b}' = [\mathbf{b}_1' \quad \mathbf{b}_2']$, gives

$$R(\mathbf{b}_1, \mathbf{b}_2) = \mathbf{y}'[\mathbf{X}_1 \quad \mathbf{X}_2]\begin{bmatrix} \mathbf{X}_1'\mathbf{X}_1 & \mathbf{X}_1'\mathbf{X}_2 \\ \mathbf{X}_2'\mathbf{X}_1 & \mathbf{X}_2'\mathbf{X}_2 \end{bmatrix}^{-}\begin{bmatrix} \mathbf{X}_1' \\ \mathbf{X}_2' \end{bmatrix}\mathbf{y} \tag{13}$$

and

$$R(\mathbf{b}_2) = \mathbf{y}'\mathbf{X}_2(\mathbf{X}_2'\mathbf{X}_2)^{-}\mathbf{X}_2'\mathbf{y} .$$

$R(\mathbf{b}_1|\mathbf{b}_2)$ is still defined as in (4) but simplification to something like (9) now depends upon which of the many available generalized inverses is used in (13), in contrast to the unique regular inverse available in (8) in the full rank model case.

Marsaglia and Styan [1974] discuss generalized inverses of partitioned matrices at length. They show that

$$\begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & (\mathbf{X}_2'\mathbf{X}_2)^{-} \end{bmatrix} + \begin{bmatrix} \mathbf{I} \\ -(\mathbf{X}_2'\mathbf{X}_2)^{-}\mathbf{X}_2'\mathbf{X}_1 \end{bmatrix}\mathbf{W}^{-}[\mathbf{I} \quad -\mathbf{X}_1'\mathbf{X}_2(\mathbf{X}_2'\mathbf{X}_2)^{-}] \tag{14}$$

for

$$\mathbf{W} = \mathbf{X}_1'[\mathbf{I} - \mathbf{X}_2(\mathbf{X}_2'\mathbf{X}_2)^{-}\mathbf{X}_2']\mathbf{X}_1 = \mathbf{X}_1'\mathbf{X}_1 - \mathbf{X}_1'\mathbf{X}_2(\mathbf{X}_2'\mathbf{X}_2)^{-}\mathbf{X}_2'\mathbf{X}_1$$

is always a generalized inverse of

$$\begin{bmatrix} \mathbf{X}_1'\mathbf{X}_1 & \mathbf{X}_1'\mathbf{X}_2 \\ \mathbf{X}_2'\mathbf{X}_1 & \mathbf{X}_2'\mathbf{X}_2 \end{bmatrix} .$$

Hence solutions of the normal equations are

$$\begin{bmatrix} \mathbf{b}_1^o \\ \mathbf{b}_2^o \end{bmatrix} = (\mathbf{X}'\mathbf{X})^- \begin{bmatrix} \mathbf{X}_1'\mathbf{y} \\ \mathbf{X}_2'\mathbf{y} \end{bmatrix} = \begin{bmatrix} \mathbf{W}^-\mathbf{X}_1'[\mathbf{I} - \mathbf{X}_2(\mathbf{X}_2'\mathbf{X}_2)^-\mathbf{X}_2]\mathbf{y} \\ (\mathbf{X}_2'\mathbf{X}_2)^-\mathbf{X}_2'(\mathbf{y} \quad -\mathbf{X}_1\mathbf{b}_1^o) \end{bmatrix} , \tag{15}$$

and using (14) in (13) therefore gives

$$R(\mathbf{b}_1|\mathbf{b}_2) = \mathbf{y}'[\mathbf{X}_1 \quad \mathbf{X}_2]\left\{\begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & (\mathbf{X}_2'\mathbf{X}_2)^- \end{bmatrix} + \begin{bmatrix} \mathbf{I} \\ -(\mathbf{X}_2'\mathbf{X}_2)^-\mathbf{X}_2'\mathbf{X}_1 \end{bmatrix} \mathbf{W}^-[\mathbf{I} \quad -\mathbf{X}_1'\mathbf{X}_2(\mathbf{X}_2'\mathbf{X}_2)^-]\right\}\begin{bmatrix} \mathbf{X}_1'\mathbf{y} \\ \mathbf{X}_2'\mathbf{y} \end{bmatrix}$$

$$- \mathbf{y}'\mathbf{X}_2(\mathbf{X}_2'\mathbf{X}_2)^-\mathbf{X}_2'\mathbf{y}$$

$$= \mathbf{y}'[\mathbf{I} - \mathbf{X}_2(\mathbf{X}_2'\mathbf{X}_2)^-\mathbf{X}_2']\mathbf{X}_1\mathbf{W}^-\mathbf{X}_1[\mathbf{I} - \mathbf{X}_2(\mathbf{X}_2'\mathbf{X}_2)^-\mathbf{X}_2']\mathbf{y} . \tag{16}$$

Furthermore, if we confine ourselves to symmetric reflexive generalized inverses $\mathbf{W}^-$ of $\mathbf{W}$, then from $\mathbf{b}_1^o$ of (15) we have (16) as

$$R(\mathbf{b}_1|\mathbf{b}_2) = \mathbf{b}_1^{o'}\mathbf{W}\mathbf{b}_1^o . \tag{17}$$

Result (17), which is illustrated in equation (A9) of the Appendix, is a generalization of $\hat{\mathbf{b}}_1'\mathbf{T}_{11}^{-1}\hat{\mathbf{b}}_1$ in the full rank case. Note, though, that (17) has to be used with much more care and circumspection than does $\hat{\mathbf{b}}_1\mathbf{T}_{11}^{-1}\hat{\mathbf{b}}_1$. The terms $\hat{\mathbf{b}}_1$ and $\mathbf{T}_{11}^{-1}$ are unique (given $\mathbf{X}_1$, $\mathbf{X}_2$ and $\mathbf{y}$). But $\mathbf{b}_1^o$ and $\mathbf{W}$ are not: there are many possible values for them. Equation (17) holds true only when it is calculated using a $\mathbf{W}$ that corresponds to $\mathbf{b}_1$ and as specified for (14); and, the $\mathbf{W}^-$ in (15) must be symmetric and reflexive. Under these conditions, (17) holds for the $\mathbf{b}_1$ corresponding to $\mathbf{W}$, but not for a $\mathbf{b}_1$ that is any other subset of $\underset{\sim}{b}$. When (17) is used correctly, it could be called the "generalized invert part of the inverse" rule.

4. Full Rank Calculations for Non-Full Rank Models

Searle [1971, Secs. 5.6 and 5.7] distinguishes between imposing, on the one hand, constraints on the elements of a solution vector (in order to obtain a solution) and, on the other hand, restrictions on a model when and if such restrictions are appropriate. The constraints on solutions that he proposes as being the easiest computationally are those of putting $p - r$ elements of the solution vector equal to zero (Searle [1971, Sec. 5.7b]), where p is the order of $\underset{\sim}{b}$ and r is the rank of $\underset{\sim}{X}$, and the $p - r$ elements are so chosen that the normal equations reduce to equations of full rank. In using this procedure there is no thought or suggestion of using restrictions on the model comparable to these constraints. The constraints are used solely as a means of getting a solution and, having once obtained a solution, it can be used for calculations pertaining to the over-parameterized, non-full rank model.

In contrast, when constraints of the form $\Sigma\alpha_i^o = 0$ and $\Sigma\beta_j^o = 0$ are used to obtain a solution vector of normal equations, the interpretation of ensuing calculations is often accompanied by the (perhaps tacit) assumption that comparable restrictions $\Sigma\alpha_i = 0$ and $\Sigma\beta_j = 0$ apply to the model. Indeed, the genuine usefulness of these constraints and restrictions in balanced data models is, in the eyes of many users of computer packages, just cause for using them with unbalanced data. It is therefore of interest to see the consequences of this, particularly in combination with using the $Q_1 = \hat{\underset{\sim}{b}}_1'\underset{\sim}{T}_{11}^{-1}\hat{\underset{\sim}{b}}_1$ calculation applied to the full rank model that results from applying these restrictions to the non-full rank model. The important consequence is that $Q_1 = \hat{\underset{\sim}{b}}_1'\underset{\sim}{T}_{11}^{-1}\hat{\underset{\sim}{b}}_1$ has different interpretations in different non-full rank models; and, in the case of data with empty cells, it has different interpretations for different patterns of empty cells, in the same non-full rank models. This is illustrated at the end of Sec. 10.2.

5. The 2-way Crossed Classification

We confine attention to the 2-way crossed classification model, first without interaction and then with interaction.

5.1. Without interaction

The model is

$$y_{ijk} = \mu + \alpha_i + \beta_j + e_{ijk} \tag{18}$$

where μ is a general mean, α_i is the effect due to the i'th level of the A-factor, β_j is that due to the j'th level of the B-factor, and e_{ijk} is the residual error term; we let $i = 1, \cdots, a$, $j = 1, \cdots, b$ and $k = 1, \cdots, n_{ij}$ with $n_{ij} \geq 0$ for all i and j.

In this model we know how to compute $R(\alpha|\mu,\beta)$: for example, in Searle [1971] at equation (39) on p. 273 and/or equation (69) on p. 297. And we also know (ibid., p. 282) that

$$\text{using } R(\alpha|\mu,\beta) \text{ tests } H: \alpha_i\text{'s all equal .} \tag{19}$$

Suppose, though, that the model is changed to one involving parameters to be denoted by $\dot{\mu}$, $\dot{\alpha}_i$ and $\dot{\beta}_j$, changed in such a manner that these new parameters are defined in terms of those of (18) in the following way:

$$\dot{\mu} = \mu + \bar{\alpha}_. + \bar{\beta}_. , \quad \dot{\alpha}_i = \alpha_i - \bar{\alpha}_. , \quad \text{and} \quad \dot{\beta}_j = \beta_j - \bar{\beta}_. . \tag{20}$$

where $\bar{\alpha}_.$ and $\bar{\beta}_.$ are averages, in the usual manner; i.e., $\bar{\alpha}_. = \sum_{i=1}^{a} \alpha_i/a$ and $\bar{\beta}_. = \sum_{j=1}^{b} \beta_j/b$. Then in the new model

$$\sum_{i=1}^{a} \dot{\alpha}_i \equiv 0 \quad \text{and} \quad \sum_{j=1}^{b} \dot{\beta}_j \equiv 0 \ . \tag{21}$$

We refer to (21) as the Σ-restrictions, and correspondingly to a model that includes them as the Σ-restricted model; (21) is illustrated in (A10).

The Σ-restrictions enable the model to be written as a full rank model in terms of $\dot{\mu}$, and the elements of $\underset{\sim}{b}_{\dot{\alpha}} \equiv \{\dot{\alpha}_i\}$ for $i = 1, \cdots, a-1$, and $\underset{\sim}{b}_{\dot{\beta}} \equiv \{\dot{\beta}_j\}$ for $j = 1, \cdots, b-1$. Then, because the model is full rank, we define the symbol

$$Q_{\dot{\alpha}} = \hat{\underset{\sim}{b}}'_{\dot{\alpha}} \underset{\sim}{T}^{-1}_{\dot{\alpha}\dot{\alpha}} \hat{\underset{\sim}{b}}_{\dot{\alpha}} \tag{22}$$

analogous to Q_1 of (7), and so use (7) to state

$$\text{using } Q_{\dot{\alpha}} \text{ tests } H: \underset{\sim}{b}_{\dot{\alpha}} = \underset{\sim}{0} \ . \tag{23}$$

Notice that the hypothesis in (23) is testable, not only because the elements $\dot{\alpha}_i$ for $i = 1, \cdots, a-1$ of $\underset{\sim}{b}_{\dot{\alpha}}$ are estimable (because they are parameters of a full rank model), but also because in (20) they are estimable functions of the parameters in the non-full rank model. Equation (A13) illustrates (22).

The hypotheses in (19) and (23) are the same. This is so because $\dot{\alpha}_i = 0$ for $i = 1, \cdots, a-1$ in (23), together with $\dot{\alpha}_a = -\sum_{i=1}^{a-1} \dot{\alpha}_i$ from (21), implies $\alpha_i = \bar{\alpha}_.$ in (20), i.e., implies that the α_i's are all equal, which is the hypothesis in (19). Therefore, because the hypotheses of (19) and (23) are the same:

$$\text{for the no interaction model, } Q_{\dot{\alpha}} = R(\alpha|\mu, \beta) \ . \tag{24}$$

5.2. With interaction, all cells filled

The model is the same as (18) but with inclusion of an interaction term γ_{ij}:

$$y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_{ij} + e_{ijk} \ . \tag{25}$$

To begin with, we restrict ourselves to the case of all cells filled, i.e., $n_{ij} > 0$ for all i and j. Searle [1971, p. 371, 2nd and subsequent printings] indicates that in this case

$$\text{using } SSA_w \text{ tests } H: \alpha_i + \bar{\gamma}_{i.} \text{ equal for all } i\ , \tag{26}$$

where SSA_w is the sum of squares for the A-effect in the weighted squares of means analysis. This is defined as

$$SSA_w = \sum_{i=1}^{a} w_i(\bar{x}_{i.} - \bar{x}_{[1]})^2 \tag{27}$$

with

$$x_{ij} = \bar{y}_{ij.}, \qquad 1/w_i = \sum_{j=1}^{b}(1/n_{ij})/b^2 \quad \text{and} \quad \bar{x}_{[1]} = \sum_{i=1}^{a} w_i\bar{x}_{i.}/\sum_{i=1}^{a} w_i\ . \tag{28}$$

The values of SSA_w and SSB_w for the numerical example are given in (A1).

Result (26) is established in Searle [1971] by deriving $E(SSA_w)$ and observing that it is zero when the $(\alpha_i + \bar{\gamma}_{i.})$ are all equal, for $i = 1, \cdots, a$. It can also be established from (11) and (12), as is done in Appendix B.

Suppose we now change the model by defining

$$\begin{aligned} \dot{\mu} &= \mu + \bar{\alpha}_. + \bar{\beta}_. + \bar{\gamma}_{..}, & \dot{\gamma}_{ij} &= \gamma_{ij} - \bar{\gamma}_{i.} - \bar{\gamma}_{.j} + \bar{\gamma}_{..}\ , \\ \dot{\alpha}_i &= \alpha_i - \bar{\alpha}_. + \bar{\gamma}_{i.} - \bar{\gamma}_{..}, \quad \text{and} & \dot{\beta}_j &= \beta_j - \bar{\beta}_. + \bar{\gamma}_{.j} - \bar{\gamma}_{..}\ . \end{aligned} \tag{29}$$

Then in the new model

$$\sum_{i=1}^{a}\dot{\alpha}_i \equiv 0, \quad \sum_{j=1}^{b}\dot{\beta}_j \equiv 0, \quad \sum_{j=1}^{b}\dot{\gamma}_{ij} \equiv 0 \ \ \forall i \quad \text{and} \quad \sum_{i=1}^{a}\dot{\gamma}_{ij} \equiv 0 \ \ \forall j\ .$$

These identities enable the model to be written as a full rank model in terms of $\underset{\sim}{b}_{\dot{\alpha}} = \{\dot{\alpha}_i\}$, of $\underset{\sim}{b}_{\dot{\beta}} = \{\dot{\beta}_j\}$ and of $\dot{\gamma}_{ij}$ for $i = 1, \cdots, a-1$ and $j = 1, \cdots, b-1$. And, because the model <u>is</u> full rank, we can use (7) to state

$$\text{using } Q_{\dot\alpha} \text{ tests } H : \underset{\sim}{b}_{\dot\alpha} = \underset{\sim}{0} \ . \tag{30}$$

But since the hypothesis here is $H : \dot\alpha_i = 0$ for $i = 1, \cdots, a-1$, this, by the definition of $\dot\alpha_i$ in (29), is identical to $H : \alpha_i + \bar\gamma_{i.}$ equal for all i, the hypothesis of (24). Hence

$$\text{in the interaction model, with all cells filled, } Q_{\dot\alpha} = SSA_w \ . \tag{31}$$

Note, of course, that if we choose to define the γ_{ij}-terms such that $\sum_{j=1}^{b} \gamma_{ij} = 0$, i.e., $\gamma_{i.} = 0$, then the sum of squares computed by either of the equivalent forms in (31) is testing $H : \alpha_i$ equal for all i.

Note also, that (31) differs from (24): $Q_{\dot\alpha}$ does not have the same interpretation in the with-interaction model (all cells filled) as in the no-interaction model.

5.3. With interaction, and some cells empty

The model in this case is exactly the same as (25) except that $n_{ij} = 0$ is now permitted, i.e., $n_{ij} \geq 0$, in contrast to $n_{ij} > 0$ of the all cells filled case.

The difficulty with the some-cells-empty case is that in terms of neither the over-parameterized model nor the full rank Σ-restricted model is it usually possible to test meaningful and/or interesting hypotheses about parameters that appear to represent either rows or columns. For example, we know that neither $R(\alpha|\mu)$ nor $R(\alpha|\mu,\beta)$ can be used to test such hypotheses, e.g., Searle [1971, equations (100) and (104), pp. 307-308]. Nor does SSA_w, which in fact is undefined, because not all cells are filled; nor does $Q_{\dot\alpha}$, as we now illustrate.

The general case is sufficiently difficult to deal with, that we confine attention to two examples.

Example 1

Consider a simple case of 2 rows and 3 columns with one empty cell, such that the pattern of filled cells is

Example 1

✓	✓	✓
✓	✓	

where ✓ indicates the presence of data. For the model (25) the cell means and their marginal averages are as shown in Table 1.

Table 1. Cell Means for the Unrestricted Model

$\mu+\alpha_1+\beta_1+\gamma_{11}$	$\mu+\alpha_1+\beta_2+\gamma_{12}$	$\mu+\alpha_1+\beta_3+\gamma_{13}$
$\mu+\alpha_2+\beta_1+\gamma_{21}$	$\mu+\alpha_2+\beta_2+\gamma_{22}$	(no data)
$\mu+\frac{1}{2}(\alpha_1+\alpha_2)+\beta_1+\frac{1}{2}(\gamma_{11}+\gamma_{21})$	$\mu+\frac{1}{2}(\alpha_1+\alpha_2)+\beta_2+\frac{1}{2}(\gamma_{12}+\gamma_{22})$	$\mu+\alpha_1+\beta_3+\gamma_{13}$

Row means	$\mu+\alpha_1+\frac{1}{3}(\beta_1+\beta_2+\beta_3)+\frac{1}{3}(\gamma_{11}+\gamma_{12}+\gamma_{13})$
	$\mu+\alpha_2+\frac{1}{2}(\beta_1+\beta_2)+\frac{1}{2}(\gamma_{21}+\gamma_{22})$

Suppose, as before, that we now change the model to be one with parameters satisfying the Σ-restrictions:

$$\begin{aligned}
\dot\alpha_1+\dot\alpha_2 &= 0\,, &\text{i.e.,}\quad \dot\alpha_2 &= -\dot\alpha_1\\
\dot\beta_1+\dot\beta_2+\dot\beta_3 &= 0 & \dot\beta_3 &= -\dot\beta_1-\dot\beta_2\\
\dot\gamma_{11}+\dot\gamma_{21} &= 0 & \dot\gamma_{21} &= -\dot\gamma_{11}\\
\dot\gamma_{21}+\dot\gamma_{22} &= 0 & \dot\gamma_{22} &= -\dot\gamma_{21} = \dot\gamma_{11}\\
\dot\gamma_{12}+\dot\gamma_{22} &= 0 & \dot\gamma_{12} &= -\dot\gamma_{22} = -\dot\gamma_{11}\\
\dot\gamma_{11}+\dot\gamma_{12}+\dot\gamma_{13} &= 0 & \dot\gamma_{13} &= -\dot\gamma_{11}-\dot\gamma_{12} = -\dot\gamma_{11}+\dot\gamma_{11} = 0\,.
\end{aligned}\tag{32}$$

The right-hand half of (32) shows how these restrictions are used to reduce the number of parameters in the model to precisely the number of cell means available in the data, in this case 5. And notice, in the last such equation, that even though cell 1,3 contains data, the $\dot{\gamma}$-term corresponding to it, $\dot{\gamma}_{13}$, is defined as zero. This is a direct consequence of the Σ-restrictions being used in combination with empty cells. The cell means and marginal averages of them, in a table similar to Table 1, become as shown in Table 2.

Table 2. Cell Means in a Model with Σ-restrictions

$\dot{\mu} + \dot{\alpha}_1 + \dot{\beta}_1 + \dot{\gamma}_{11}$	$\dot{\mu} + \dot{\alpha}_1 + \dot{\beta}_2 - \dot{\gamma}_{11}$	$\dot{\mu} + \dot{\alpha}_1 - \dot{\beta}_1 - \dot{\beta}_2$	$\dot{\mu} + \dot{\alpha}_1$
$\dot{\mu} - \dot{\alpha}_1 + \dot{\beta}_1 - \dot{\gamma}_{11}$	$\dot{\mu} - \dot{\alpha}_1 + \dot{\beta}_2 + \dot{\gamma}_{11}$	(no data)	$\dot{\mu} - \dot{\alpha}_1 + \frac{1}{2}(\dot{\beta}_1 + \dot{\beta}_2)$
$\dot{\mu} + \dot{\beta}_1$	$\dot{\mu} + \dot{\beta}_2$	$\dot{\mu} + \dot{\alpha}_1 - \dot{\beta}_1 - \dot{\beta}_2$	

The relationship of the parameters in Table 2 to those of Table 1 is obtained by equating corresponding cell means in the two tables. This leads to the following expressions:

$$\begin{aligned}
\dot{\mu} &= \mu + \tfrac{1}{2}(\alpha_1+\alpha_2) + \tfrac{1}{3}(\beta_1+\beta_2+\beta_3) + \tfrac{1}{12}(\gamma_{11}+\gamma_{12}) + \tfrac{1}{3}\gamma_{13} + \tfrac{1}{4}(\gamma_{21}+\gamma_{22}) \\
\dot{\alpha}_1 &= \tfrac{1}{2}(\alpha_1-\alpha_2) + \tfrac{1}{4}(\gamma_{11}+\gamma_{12}-\gamma_{21}-\gamma_{22}) \\
\dot{\beta}_1 &= \beta_1 - \tfrac{1}{3}(\beta_1+\beta_2+\beta_3) + \tfrac{5}{12}\gamma_{11} - \tfrac{1}{12}\gamma_{12} - \tfrac{1}{3}\gamma_{13} + \tfrac{1}{4}\gamma_{21} - \tfrac{1}{4}\gamma_{22} \qquad (33) \\
\dot{\beta}_2 &= \beta_2 - \tfrac{1}{3}(\beta_1+\beta_2+\beta_3) - \tfrac{1}{12}\gamma_{11} + \tfrac{5}{12}\gamma_{12} - \tfrac{1}{3}\gamma_{13} - \tfrac{1}{4}\gamma_{21} + \tfrac{1}{4}\gamma_{22} \\
\dot{\gamma}_{11} &= \tfrac{1}{4}(\gamma_{11}-\gamma_{12}-\gamma_{21}+\gamma_{22})
\end{aligned}$$

Cell means of Table 1 are estimable; and equations (33) come from equating cell means of Table 2 to those of Table 1. Therefore the parameters of the new model (which are estimable because that model is of full rank) are also estimable functions of the parameters of the old model; i.e., the right-hand sides of (33) are estimable functions of the over-parameterized model (25).

Now consider a statement similar to (7) and (23):

$$\text{using } Q_{\dot{\beta}} \text{ tests } H : \underset{\sim}{b}_{\dot{\beta}} = \underset{\sim}{0} \;, \tag{34}$$

i.e., $Q_{\dot{\beta}}$ tests $H : \dot{\beta}_1 = \dot{\beta}_2 = 0$ (and also $\dot{\beta}_3 = 0$ because $\dot{\beta}_3 = -\dot{\beta}_1 - \dot{\beta}_2$). Whereas this hypothesis is easily understood in terms of the symbols $\dot{\beta}_1$, $\dot{\beta}_2$ (and $\dot{\beta}_3$), the meaning of those symbols is not so easily understood. For example, in Table 2, $\dot{\mu} + \dot{\beta}_1$ and $\dot{\mu} + \dot{\beta}_2$ are the means for columns 1 and 2, and we might well expect $\dot{\mu} - \dot{\beta}_1 - \dot{\beta}_2$ to be the mean for column 3; but it is not. Nor is the meaning of (34) very clear when thought of in terms of the estimable functions for $\dot{\beta}_1$ and $\dot{\beta}_2$ given in (33).

Similarly,

$$\text{using } Q_{\dot{\alpha}} \text{ tests } H : \dot{\alpha}_1 = 0 \text{ (and also } \dot{\alpha}_2 = 0) \;, \tag{35}$$

and in terms of (33) this hypothesis is

$$H : \alpha_1 + \tfrac{1}{2}(\gamma_{11} + \gamma_{12}) = \alpha_2 + \tfrac{1}{2}(\gamma_{21} + \gamma_{22}) \;. \tag{36}$$

This is the hypothesis that rows, in the presence of averaged interactions, are equal over the first two columns. In some situations this may well be a useful hypothesis; but what is important is that the hypothesis in (35) is equivalent to that in (36) and is not, in any general sense, a hypothesis of equality of the rows over all three columns, as might be inferred by the symbol $Q_{\dot{\alpha}}$ and a knowledge of the full rank statement (7).

Example 2

The functions on the right-hand sides of (33) become even more complicated when there are several, or many, empty cells. Consider the data pattern

Example 2

✓	✓	✓
✓	✓	
✓		✓

The table of cell means analogous to Table 2 is Table 3.

Table 3. Cell Means in a Model with Σ-restrictions

$\dot\mu+\dot\alpha_1+\dot\beta_1+\dot\gamma_{11}$	$\dot\mu+\dot\alpha_1+\dot\beta_2+\dot\gamma_{12}$	$\dot\mu+\dot\alpha_1-\dot\beta_1-\dot\beta_2-\dot\gamma_{11}-\dot\gamma_{12}$	$\dot\mu+\dot\alpha_1$
$\dot\mu+\dot\alpha_2+\dot\beta_1+\dot\gamma_{12}$	$\dot\mu+\dot\alpha_2+\dot\beta_2-\dot\gamma_{12}$	(no data)	$\dot\mu+\dot\alpha_2+\frac{1}{2}(\dot\beta_1+\dot\beta_2)$
$\dot\mu-\dot\alpha_1-\dot\alpha_2+\dot\beta_1-\dot\gamma_{11}-\dot\gamma_{12}$	(no data)	$\dot\mu-\dot\alpha_1-\dot\alpha_2-\dot\beta_1-\dot\beta_2+\dot\gamma_{11}+\dot\gamma_{12}$	$\dot\mu-\dot\alpha_1-\dot\alpha_2-\frac{1}{2}\dot\beta_2$
$\dot\mu+\dot\beta_1$	$\dot\mu+\frac{1}{2}(\dot\alpha_1+\dot\alpha_2)+\dot\beta_2$	$\dot\mu-\frac{1}{2}\dot\alpha_2-(\dot\beta_1+\dot\beta_2)$	

Equating the cell means of Table 3 to those of a table analogous to Table 2 yields the equations comparable to (33) as follows:

$$
\begin{aligned}
\dot\mu &= \mu + \tfrac{1}{3}(\alpha_1+\alpha_2+\alpha_3) + \tfrac{1}{3}(\beta_1+\beta_2+\beta_3) + \tfrac{1}{15}(\gamma_{11}+2\gamma_{12}+2\gamma_{13}+2\gamma_{21}+3\gamma_{22}+2\gamma_{31}+3\gamma_{33})\\
\dot\alpha_1 &= \alpha_1 - \tfrac{1}{3}(\alpha_1+\alpha_2+\alpha_3) + \tfrac{1}{15}(4\gamma_{11}+3\gamma_{12}+3\gamma_{13}-2\gamma_{21}-3\gamma_{22}-2\gamma_{31}-3\gamma_{33})\\
\dot\alpha_2 &= \alpha_2 - \tfrac{1}{3}(\alpha_1+\alpha_2+\alpha_3) + \tfrac{1}{15}(-2\gamma_{11}-4\gamma_{12}+\gamma_{13}+6\gamma_{21}+4\gamma_{22}-4\gamma_{31}-\gamma_{33})\\
\dot\beta_1 &= \beta_1 - \tfrac{1}{3}(\beta_1+\beta_2+\beta_3) + \tfrac{1}{15}(4\gamma_{11}-2\gamma_{12}-2\gamma_{13}+3\gamma_{21}-3\gamma_{22}+3\gamma_{31}-3\gamma_{33}) \qquad (37)\\
\dot\beta_2 &= \beta_2 - \tfrac{1}{3}(\beta_1+\beta_2+\beta_3) + \tfrac{1}{15}(-2\gamma_{11}+6\gamma_{12}-4\gamma_{13}-4\gamma_{21}+4\gamma_{22}+\gamma_{31}-\gamma_{33})\\
\dot\gamma_{11} &= + \tfrac{1}{15}(6\gamma_{11}-3\gamma_{12}-3\gamma_{13}-3\gamma_{21}+3\gamma_{22}-3\gamma_{31}+3\gamma_{33})\\
\dot\gamma_{12} &= + \tfrac{1}{15}(-3\gamma_{11}+4\gamma_{12}-\gamma_{13}+4\gamma_{21}-4\gamma_{22}-\gamma_{31}+\gamma_{33})
\end{aligned}
$$

Then we can say

$$\text{using } Q_{\dot{\alpha}} \text{ tests } H : \dot{\alpha}_i = 0 \text{ for } i = 1,\ 2\,.$$

But interpretation of this hypothesis is difficult. Symbolically it means testing $\dot{\alpha}_1 = 0 = \dot{\alpha}_2$ $(= \dot{\alpha}_3)$. But in terms of Table 3, although $\dot{\mu} + \dot{\alpha}_1$ is the mean for row 1, $\dot{\mu} + \dot{\alpha}_2$ is not the mean for row 2; and in terms of (37), $\dot{\alpha}_1$ and $\dot{\alpha}_2$ involve functions of the α's that are easily interpreted and functions of the γ's that are not.

5.4. A general relationship between models

The specific form of relationships such as (33) and (37) depends upon the unrestricted (over-parameterized) model itself, upon the restrictions being used (in this case the Σ-restrictions) and upon the pattern of empty cells. Nevertheless, the relationships between models can be specified as follows for the general case.

Let $\underset{\sim}{b}$ be the vector of parameters in the unrestricted model, and let $\underset{\sim}{b}_s$ (of order s, the number of filled cells in the data) be the vector of parameters in the restricted, full rank model. If $\underset{\sim}{X}$ and $\underset{\sim}{X}_s$ are correspondingly defined by $E(\underset{\sim}{y}) = \underset{\sim}{X}\underset{\sim}{b}$ and $E(\underset{\sim}{y}) = \underset{\sim}{X}_s\underset{\sim}{b}_s$, then define $\underset{\sim}{W}$ and $\underset{\sim}{W}_s$ as matrices of the s distinctly different rows of $\underset{\sim}{X}$ and $\underset{\sim}{X}_s$, respectively. Then $\underset{\sim}{b}$ and $\underset{\sim}{b}_s$ are related by equating cell means $\underset{\sim}{W}_s\underset{\sim}{b}_s = \underset{\sim}{W}\underset{\sim}{b}$. When all possible interactions exist $\underset{\sim}{W}_s^{-1}$ exists, and so

$$\underset{\sim}{b}_s = \underset{\sim}{W}_s^{-1}\underset{\sim}{W}\underset{\sim}{b}\,. \tag{38}$$

More generally, $\underset{\sim}{b}_s = E(\hat{\underset{\sim}{b}}_s)$ becomes $\underset{\sim}{b}_s = (\underset{\sim}{X}_s'\underset{\sim}{X}_s)^{-1}\underset{\sim}{X}_s'\underset{\sim}{X}\underset{\sim}{b}$. If $(\underset{\sim}{X}'\underset{\sim}{X})^-$ is the generalized inverse of $\underset{\sim}{X}'\underset{\sim}{X}$ corresponding to the restrictions, with $\underset{\sim}{b}^o = (\underset{\sim}{X}'\underset{\sim}{X})^-\underset{\sim}{X}'\underset{\sim}{y}$, then since $\hat{\underset{\sim}{b}}_s$ is a subset of $\underset{\sim}{b}^o$, i.e., $\hat{\underset{\sim}{b}}_s = \underset{\sim}{S}\underset{\sim}{b}^o$ (so defining $\underset{\sim}{S}$), we have $\underset{\sim}{b}_s = E(\hat{\underset{\sim}{b}}_s) = \underset{\sim}{S}E(\underset{\sim}{b}^o) = \underset{\sim}{S}\underset{\sim}{H}\underset{\sim}{b}$, on defining $\underset{\sim}{H} = (\underset{\sim}{X}'\underset{\sim}{X})^-\underset{\sim}{X}'\underset{\sim}{X}$.

In Example 1, the parameters in Tables 1 and 2 are $\underset{\sim}{b}$ and $\underset{\sim}{b}_s$, respectively. Then (38) is

$$\begin{bmatrix} \dot{\mu} \\ \dot{\alpha}_1 \\ \dot{\beta}_1 \\ \dot{\beta}_2 \\ \dot{\gamma}_{11} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 & -1 \\ 1 & 1 & -1 & -1 & 0 \\ 1 & -1 & 1 & 0 & -1 \\ 1 & -1 & 0 & 1 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 1 & 1 & . & 1 & . & . & 1 & . & . & . & . \\ 1 & 1 & . & . & 1 & . & . & 1 & . & . & . \\ 1 & 1 & . & . & . & 1 & . & . & 1 & . & . \\ 1 & . & 1 & 1 & . & . & . & . & . & 1 & . \\ 1 & . & 1 & . & 1 & . & . & . & . & . & 1 \end{bmatrix} \begin{bmatrix} \mu \\ \alpha_1 \\ \alpha_2 \\ \beta_1 \\ \beta_2 \\ \beta_3 \\ \gamma_{11} \\ \gamma_{12} \\ \gamma_{13} \\ \gamma_{21} \\ \gamma_{22} \end{bmatrix},$$

which reduces to (33).

6. μ_{ij}-Models

There is growing acceptance of the fact that models of the form (25) should be replaced by models like

$$y_{ijk} = \mu_{ij} + e_{ijk}\,, \tag{39}$$

especially when there are cells with no data in them. When such a model is also to include the specification of no interactions between rows and columns, this can be expressed in the form

$$\underset{\sim}{P}'\underset{\sim}{\mu} = \underset{\sim}{0} \tag{40}$$

as an equation to accompany (39) as part of the model, $\underset{\sim}{\mu}$ being the vector of μ_{ij}'s in (39) and $\underset{\sim}{P}'$ being a known matrix. Estimation for the model (39) is

$$\hat{\underset{\sim}{\mu}} = \bar{\underset{\sim}{y}} = \{\bar{y}_{ij.}\}\,, \tag{41}$$

the vector of cell means in the data; and for the model (39) and (40), estimation is

$$\hat{\underset{\sim}{\mu}}_r = \bar{\underset{\sim}{y}} - \underset{\sim}{D}\underset{\sim}{P}(\underset{\sim}{P}'\underset{\sim}{D}\underset{\sim}{P})^{-1}\underset{\sim}{P}'\bar{\underset{\sim}{y}} \tag{42}$$

where

$$\underset{\sim}{D} = \underset{\sim}{D}\{1/n_{ij}\} \tag{43}$$

is the diagonal matrix of terms $1/n_{ij}$ for n_{ij} being the number of observations in cell (i,j) that has data. [Equation (42) is a special case of, for example, Searle [1971], equation (97), p. 206.]

Examples

For Example 1 the no-interaction model would be

$$y_{ijk} = \mu_{ij} + e_{ijk}, \quad \text{and} \quad \mu_{11} - \mu_{12} - \mu_{21} + \mu_{22} = 0 .$$

In this case $\underset{\sim}{P}' = [1 \ \ -1 \ \ 0 \ \ -1 \ \ 1]$ and from (42)

$$\hat{\underset{\sim}{\mu}}_r = \underset{\sim}{\bar{y}} - \underset{\sim}{DP}(1/n_{11} + 1/n_{12} + 1/n_{21} + 1/n_{22})^{-1}(\bar{y}_{11} - \bar{y}_{12} - \bar{y}_{21} + \bar{y}_{22})$$

$$= \begin{bmatrix} \bar{y}_{11\cdot} - \lambda/n_{11} \\ \bar{y}_{12\cdot} + \lambda/n_{12} \\ \bar{y}_{13\cdot} \\ \bar{y}_{21\cdot} + \lambda/n_{21} \\ \bar{y}_{22\cdot} - \lambda/n_{22} \end{bmatrix} \quad \text{for} \quad \lambda = \frac{\bar{y}_{11.} - \bar{y}_{12.} - \bar{y}_{21.} + \bar{y}_{22.}}{\dfrac{1}{n_{11}} + \dfrac{1}{n_{12}} + \dfrac{1}{n_{21}} + \dfrac{1}{n_{22}}} .$$

And for Example 2, the matrix $\underset{\sim}{P}'$ to be used in (42) is

$$\underset{\sim}{P}' = \begin{bmatrix} 1 & -1 & 0 & -1 & 1 & 0 & 0 \\ 1 & 0 & -1 & 0 & 0 & -1 & 1 \end{bmatrix} .$$

The great utility of models like (39), or (39) and (40) together, is that all of them are of full rank and in all of them every μ_{ij} (corresponding to a cell containing data) is estimable, and so is every linear function of such μ_{ij}'s.

This permits the researcher to test any linear hypothesis about the cell means that interests him. True it is, that these models do not have built-in definitions of what we familiarly call row effects or column effects. Nevertheless, when data have empty cells, our apparent definitions of these effects do not always measure up to their appearances – as we have just illustrated, with simple examples. In contrast, the μ_{ij}-models enable a researcher to define a row effect as any linear combination of cell means that seems appropriate. Thus in Example 1, to estimate an effect due to row 1, that effect might be defined as $\frac{1}{3}(\mu_{11} + \mu_{12} + \mu_{13})$; but to compare rows 1 and 2, it would probably be defined as $\frac{1}{2}(\mu_{11} + \mu_{12})$. There need be no confusion in having two such definitions: the first is the effect averaged over all columns, whereas the second is the effect averaged over only those columns wherein there are also data on row 2, thus permitting a comparison of rows 1 and 2 over the same columns.

Our use in this paper for the μ_{ij}-models is that of possibly aiding interpretation of the parameters in Σ-restricted models. We do this with the help of equation (38). In $\underset{\sim\sim}{Wb}$ of that equation, the matrix of coefficients of the γ_{ij}'s (in general of the highest-order interaction effects) is an identity matrix, so that in (38) itself the matrix of coefficients of the γ_{ij}'s is $\underset{\sim}{W}_s^{-1}$. Furthermore, if in (38) $\underset{\sim}{b}$ were to be the parameter vector for the μ_{ij}-model, $\underset{\sim\sim}{Wb}$ would be $\underset{\sim\sim}{I\mu}$. Then (38) would yield $\underset{\sim}{b}_s = \underset{\sim}{W}_s^{-1}\underset{\sim}{\mu}$. Thus $\underset{\sim}{b}_s$ expressed in terms of the μ_{ij}-model is the same function of the μ_{ij}'s as is the function of γ_{ij}'s occurring in $\underset{\sim}{b}_s = \underset{\sim}{W}_s^{-1}\underset{\sim\sim}{Wb}$ of (38). Hence, for example, from (33)

$$
\begin{aligned}
\dot{\mu} &= \tfrac{1}{12}(\mu_{11} + \mu_{12}) + \tfrac{1}{3}\gamma_{13} + \tfrac{1}{4}(\mu_{21} + \mu_{22}) \\
\dot{\alpha}_1 &= \tfrac{1}{4}(\mu_{11} + \mu_{12} - \mu_{21} - \mu_{22}) \\
\dot{\beta}_1 &= \tfrac{5}{12}\mu_{11} - \tfrac{1}{12}\mu_{12} - \tfrac{1}{3}\mu_{13} + \tfrac{1}{4}\mu_{21} - \tfrac{1}{4}\mu_{22} \qquad (44) \\
\dot{\beta}_2 &= -\tfrac{1}{12}\mu_{11} + \tfrac{5}{12}\mu_{12} - \tfrac{1}{3}\mu_{13} - \tfrac{1}{4}\mu_{21} + \tfrac{1}{4}\mu_{22} \\
\dot{\gamma}_{11} &= \tfrac{1}{4}(\mu_{11} - \mu_{12} - \mu_{21} + \mu_{22}) \;.
\end{aligned}
$$

A pictorial way of representing these combinations of cell means is to express them all with a common denominator, 12 in this case, and then display the numerators of the coefficients of the μ_{ij}'s in a grid corresponding to the pattern of filled cells in the data. This is the basis of Table 4.

Table 4. Coefficients[1] of μ_{ij}'s in Equations (44), laid out in the pattern of the data:

	μ_{11} / μ_{21}	μ_{12} / μ_{22}	μ_{13}
$\dot{\mu}$	1	1	4
	3	3	
$\dot{\alpha}_1$	3	3	
	-3	-3	
$\dot{\beta}_1$	5	-1	-4
	3	-3	
$\dot{\beta}_2$	-1	5	-4
	-3	3	
$\dot{\gamma}_{11}$	3	-3	
	-3	3	

[1] Each coefficient has to be divided by 12.

In this way it is easily seen that $\dot{\alpha}_1$ is the difference between rows 1 and 2 averaged over columns 1 and 2 (as has already been observed); and $\dot{\gamma}_{11}$ is the interaction effect of rows 1 and 2 with columns 1 and 2. In contrast, useful interpretation of $\dot{\beta}_1$ and $\dot{\beta}_2$ seems to be more difficult.

Equations (37) for Example 2 can be treated similarly. The right-hand sides can be expressed as the same functions of μ_{ij}'s as are the functions of γ_{ij}'s in those right-hand sides. Then, with a common denominator 15, they can be summarized as in Table 5.

Table 5. Coefficients[1/] of μ_{ij}'s Coming from Equations (37), laid out in the pattern of the data:

	μ_{11}	μ_{12}	μ_{13}
	μ_{21}	μ_{22}	
	μ_{31}		μ_{33}
	1	2	2
$\dot{\mu}$	2	3	
	2		3
	4	3	3
$\dot{\alpha}_1$	-2	-3	
	-2		-3
	-2	-4	1
$\dot{\alpha}_2$	6	4	
	-4		-1
	4	-2	-2
$\dot{\beta}_1$	3	-3	
	3		-3
	-2	6	-4
$\dot{\beta}_2$	-4	4	
	1		-1
	6	-3	-3
$\dot{\gamma}_{11}$	-3	3	
	-3		3
	-3	4	-1
$\dot{\gamma}_{12}$	4	-4	
	-1		1

[1/] Each coefficient has to be divided by 15.

It is clear from this that $\dot{\gamma}_{11}$ is the sum of the interaction of rows 1 and 2 with columns 1 and 2, and the interaction of rows 1 and 3 with columns 1 and 3; whereas $\dot{\gamma}_{12}$ is a weighted difference of these two interactions, the first of them weighted by -4 and the second by +1. Similar description of the other terms in Table 5 does not appear to be quite so straightforward.

In looking at the details of Examples 1 and 2, one should not be led into false comfort that interpretation of parameters in a full rank Σ-restricted model is always easy or useful. It is both of these things in the case of equal-subclass-numbers data, and it is mostly that way for the case of all cells filled; but for the case of some cells empty it is generally neither easy nor useful. Furthermore, the exact interpretation depends upon just which cells are empty. Examples 1 and 2 are relatively simple cases, in terms of the pattern of empty cells, and yet the interpretation of the Σ-restricted parameters is not easy. When data are more extensive than these examples, and with more empty cells, then interpretation will be even more complicated.

7. Extending the $R(\cdot)$ Notation to Restricted Models

Consider the full rank, Σ-restricted, 2-way classification interaction model discussed in Sections 5.2 and 5.3. The reduction in sum of squares for fitting that model will be denoted at first by $R(\dot{\mu},\dot{\alpha},\dot{\beta},\dot{\gamma})$. It can be calculated from the R-algorithm in (2) applied to the normal equations of the full rank, Σ-restricted model. But these normal equations are nothing more than a full rank representation (or reparameterization) of those for the unrestricted model (25). Therefore

$$R(\dot{\mu},\dot{\alpha},\dot{\beta},\dot{\gamma}) = R(\mu,\alpha,\beta,\gamma) \ , \tag{45}$$

the reduction in sum of squares due to fitting the unrestricted model.

Now consider $R(\dot{\mu},\dot{\beta},\dot{\gamma})$ calculated from the R-algorithm applied to equations that come from the normal equations of the Σ-restricted model after amending them by deleting the $\dot{\alpha}_i$'s and the $\dot{\alpha}_i$-equations (those that come from the differentiation with respect to the $\dot{\alpha}_i$'s in the least squares process). Then, because we are dealing with a full rank model, we know that

$$\text{using } R(\dot{\mu},\dot{\alpha},\dot{\beta},\dot{\gamma}) - R(\dot{\mu},\dot{\beta},\dot{\gamma}) \text{ tests } H: \dot{\alpha}_i = 0,\ i = 1, \cdots, a-1 . \tag{46}$$

But from (7), we also know that

$$\text{using } Q_{\dot{\alpha}} \text{ tests } H: \underset{\sim}{b}_{\dot{\alpha}} = \underset{\sim}{0}, \text{ i.e., } H: \alpha_i = 0,\ i = 1, \cdots, a-1. \tag{47}$$

Therefore

$$Q_{\dot{\alpha}} = R(\dot{\mu}, \dot{\alpha}, \dot{\beta}, \dot{\gamma}) - R(\dot{\mu}, \dot{\beta}, \dot{\gamma}) , \tag{48}$$

i.e.,

$$Q_{\dot{\alpha}} = R(\dot{\alpha} | \dot{\mu}, \dot{\beta}, \dot{\gamma}) . \tag{49}$$

The $R(\cdot)$ notation in (48) and (49) involves two important differences from its customary usage in, say,

$$R(\alpha | \mu, \beta) = R(\mu, \alpha, \beta) - R(\mu, \beta) . \tag{50}$$

First, the model used in (48) is a restricted model as indicated by the dots above the parameter symbols. But the form of the restrictions is not specifically indicated. This is taken care of by introducing a subscript Σ to yield the symbol $R(\dot{\alpha} | \dot{\mu}, \dot{\beta}, \dot{\gamma})_\Sigma$. Second, whereas in (50) the term $R(\mu, \beta)$ relates to fitting the model $E(y_{ij}) = \mu + \beta_j$, in (48) the term $R(\dot{\mu}, \dot{\beta}, \dot{\gamma})$ does not relate to fitting

$$E(y_{ijk}) = \dot{\mu} + \dot{\beta}_j + \dot{\gamma}_{ij} \tag{51}$$

with Σ-restrictions pertaining to <u>just</u> that model, namely

$$\sum_{j=1}^{b} \dot{\beta}_j = 0 \quad \text{and} \quad \sum_{i=1}^{a} \dot{\gamma}_{ij} = 0 \quad \forall j . \tag{52}$$

The reduction in sum of squares for fitting (51) and (52) would be denoted $R(\dot{\mu}, \dot{\beta}, \dot{\gamma})_\Sigma$. But since the normal equations for (51) and (52) would be just a full rank representation of those for the unrestricted model

$$E(y_{ijk}) = \mu + \beta_j + \gamma_{ij} \tag{53}$$

and so would yield $R(\mu,\beta,\gamma)$ as the reduction in sum of squares, this means

$$R(\dot{\mu},\dot{\beta},\dot{\gamma})_\Sigma = R(\mu,\beta,\gamma) \ . \tag{54}$$

And therefore, because (53) is indistinguishable from the nested (hierarchical) model,

$$R(\dot{\mu},\dot{\beta},\dot{\gamma})_\Sigma = R(\mu,\beta,\gamma) = \sum_i\sum_j n_{ij}\bar{y}^2_{ij\cdot} \ . \tag{55}$$

(See, for example, Searle [1971, Sec. 6.4].) Hence, because $R(\mu,\alpha,\beta,\gamma) = \sum_i\sum_j n_{ij}\bar{y}^2_{ij\cdot}$ also,

$$R(\mu,\alpha,\beta,\gamma) - R(\dot{\mu},\dot{\beta},\dot{\gamma})_\Sigma \equiv 0 \ . \tag{56}$$

In contrast to $R(\dot{\mu},\dot{\beta},\dot{\gamma})_\Sigma$, where the Σ-restrictions on the $\dot{\beta}$'s and $\dot{\gamma}$'s are those appropriate only to the model (51), namely the restrictions (52), the term $R(\dot{\mu},\dot{\beta},\dot{\gamma})$ in (48) is calculated using the Σ-restricted model $E(y_{ijk}) = \dot{\mu} + \dot{\alpha}_i + \dot{\beta}_j + \dot{\gamma}_{ij}$ that involves the restrictions (52) <u>and</u> the restrictions

$$\sum_{i=1}^{a} \dot{\alpha}_i = 0 \quad \text{and} \quad \sum_{j=1}^{b} \dot{\gamma}_{ij} = 0 \ \ \forall i \ . \tag{57}$$

From the normal equations for this model the $\dot{\alpha}_i$'s and $\dot{\alpha}_i$-equations are deleted, and the $R(\dot{\mu},\dot{\beta},\dot{\gamma})$ for (48) is then calculated by applying the R-algorithm to the remaining equations. But these remaining equations, because they come from normal equations that utilize the $\dot{\gamma}$-restrictions of (52) as well as those of (57), will be different from the normal equations for (51) and (52). Therefore the $R(\dot{\mu},\dot{\beta},\dot{\gamma})$ term for (48) will not be the same as $R(\dot{\mu},\dot{\beta},\dot{\gamma})_\Sigma$ of (54) and (55): and to notationally take account of using both (52) and (57) we denote the term by $R^*(\dot{\mu},\dot{\beta},\dot{\gamma})_\Sigma$, and then have

$$R^*(\dot{\mu},\dot{\beta},\dot{\gamma})_\Sigma \neq R(\dot{\mu},\dot{\beta},\dot{\gamma})_\Sigma = R(\mu,\beta,\gamma) \ . \tag{58}$$

Then (49) is

$$Q_{\dot{\alpha}} \equiv R(\mu,\alpha,\beta,\gamma) - R^*(\dot{\mu},\dot{\beta},\dot{\gamma})_\Sigma \tag{59}$$

$$= R^*(\dot{\alpha}|\dot{\mu},\dot{\beta},\dot{\gamma})_\Sigma\ , \tag{60}$$

and this is not zero. Thus in the notation $R^*(\dot{\alpha}|\dot{\mu},\dot{\beta},\dot{\gamma})_\Sigma$ the overhead dots indicate a fully parameterized model, the subscript Σ indicates that the reparameterization has been achieved using the Σ-restrictions, and the superscript * indicates the Σ-restrictions of the full model are used throughout, i.e., are used for calculating $R^*(\dot{\mu},\dot{\beta},\dot{\gamma})_\Sigma$ as described following (57). What this amounts to is that in $R(\dot{\mu},\dot{\beta},\dot{\gamma})_\Sigma$ the Σ-restrictions are just those of the sub-model, involving $\dot{\mu}$, $\dot{\beta}$'s and $\dot{\gamma}$'s; whereas the asterisk in $R^*(\dot{\alpha}|\dot{\mu},\dot{\beta},\dot{\gamma})_\Sigma$ indicates that the Σ-restrictions are those of the full model involving $\dot{\mu}$, $\dot{\alpha}$'s, $\dot{\beta}$'s and $\dot{\gamma}$'s, even (especially) in the calculation of $R^*(\dot{\mu},\dot{\beta},\dot{\gamma})_\Sigma$. An illustration of (60) is given in (A15) *et seq*.

The preceding description applies whether or not all cells are filled. In the particular case that they are all filled we have, from (31)

$$\text{in the all-cells-filled case, } Q_{\dot{\alpha}} = SSA_w = R^*(\dot{\alpha}|\dot{\mu},\dot{\beta},\dot{\gamma})_\Sigma\ . \tag{61}$$

8. Other Restrictions

The notation $R^*(\dot{\alpha}|\dot{\mu},\dot{\beta},\dot{\gamma})_\Sigma$ developed in the preceding section is applicable to any full rank, restricted model using restrictions other than the Σ-restrictions. We consider, for *the all-cells-filled case only*, two further sets of restrictions.

8.1. The 0-restrictions

Constraints useful for solving normal equations in the unrestricted model are those which put some elements of the solution vector equal to zero (e.g., Searle [1971, p. 213]). Analogous restrictions on the model are called the O-restrictions.

The particular set of O-restrictions considered by Speed and Hocking [1976], which we term the O_{11}-restrictions is

$$\dot{\alpha}_1 = 0 \qquad \dot{\gamma}_{1j} = 0, \ \forall j$$
$$\dot{\beta}_1 = 0 \qquad \dot{\gamma}_{i1} = 0, \ \forall i \ .$$

A generalization of these, to be called the O_{kt}-restrictions, is

for an arbitrarily chosen k and t

$$\begin{aligned} \dot{\alpha}_k &= 0 \qquad \dot{\gamma}_{kj} = 0, \ \forall j \\ \dot{\beta}_t &= 0 \qquad \dot{\gamma}_{it} = 0, \ \forall i \ . \end{aligned} \tag{62}$$

Then Speed and Hocking indicate that

$$\text{using } R^*(\dot{\beta}|\dot{\mu},\dot{\alpha},\dot{\gamma})_{O_{kt}} \text{ tests } H: \beta_j + \gamma_{kj} \text{ all equal .} \tag{63}$$

In view of (62) this hypothesis is $H: \dot{\beta}_j = 0$, but it is, of course, a hypothesis of equality of columns tested over a specified row, namely the k'th row.

Illustration of the O-restrictions is given in Sec. 10.4.

8.2. The W-restrictions

Other restrictions sometimes employed are those that involve weighted sums of the parameters, using the n_{ij}-values as weights. We call these the W-restrictions. An example is:

$$\sum_{i=1}^{a} n_{i.}\dot{\alpha}_i = 0, \qquad \sum_{i=1}^{a} n_{ij}(\dot{\alpha}_i + \dot{\gamma}_{ij}) = 0, \ \forall j \ ,$$
$$\sum_{j=1}^{b} n_{.j}\dot{\beta}_j = 0, \qquad \sum_{j=1}^{b} n_{ij}(\dot{\beta}_j + \dot{\gamma}_{ij}) = 0, \ \forall i \ , \tag{64}$$

used by Speed _et al._ [1978]. They indicate that relationships to the classical analysis of variance are of the form

$$R^*(\dot{\alpha}|\dot{\mu},\dot{\beta},\dot{\gamma})_W = R(\alpha|\mu) \ . \tag{65}$$

The W-restrictions are illustrated in Sec. 10.5.

9. References

Gosslee, D. G. and Lucas, H. L. [1965]. Analysis of variance of disproportionate data when interaction is present. _Biometrics, 21_, 115-133.

Harvey, Walter R. [1970]. Estimation of variance and covariance components in the mixed model. _Biometrics, 26_, 485-504.

Henderson, C. R. [1959, 1969]. Design and analysis of animal husbandry experiments. Chapter 1 of _Techniques and Procedures in Animal Science Research_ (1st Ed., 1959; 2nd Ed., 1969), American Society of Animal Science.

Hocking, R. R., Hackney, O. P., and Speed, F. M. [1978]. The analysis of linear models with unbalanced data. _Contributions to Survey Sampling and Applied Statistics, Festschrift for H. O. Hartley_, Ed. H. A. David, Academic Press.

Marsaglia, George and Styan, George P. H. [1974]. Equalities and inequalities for ranks of matrices. _Linear and Multilinear Algebra, 2_, 269-292.

Searle, S. R. [1966]. _Matrix Algebra for the Biological Sciences_. Wiley, New York.

Searle, S. R. [1971]. _Linear Models_. Wiley, New York.

Searle, S. R. [1972]. Using the R( )-notation for reductions in sums of squares when fitting linear models. Paper presented at the Spring Regional Meetings of ENAR, Ames, Iowa.

Speed, F. M., Hocking, R. R., and Hackney, O. P. [1978]. Methods of analysis of linear models with unbalanced data. _J. Amer. Stat. Assoc., 73_, 105-112.

Yates, F. [1934]. The analysis of multiple classifications with unequal numbers in the different classes. _J. Amer. Stat. Assoc., 29_, 51-66.

10. Appendix A: Numerical Illustration

We use the following hypothetical data by way of illustration.

Table A1: Observations

	$j = 1$	$j = 2$	$j = 3$
$i = 1$	7,9	6	2
$i = 2$	8	4,8	12

Table A2: Totals

	$y_{ij.}$			$y_{i..}$
	16	6	2	24
	8	12	12	32
$y_{.j.}$	24	18	14	$56 = y_{...}$

Table A3: Numbers

	n_{ij}			$n_{i.}$
	2	1	1	4
	1	2	1	4
$n_{.j}$	3	3	2	$8 = N$

Table A4: Means

	$\bar{y}_{ij.}$			$\bar{y}_{i..}$
	8	6	2	6
	8	6	12	8
$\bar{y}_{.j.}$	8	6	7	$7 = \bar{y}_{...}$

10.1. The unrestricted, over-parameterized, model

Using standard procedures for calculating the classical partitionings of sums of squares (e.g., Searle [1971, Sec. 7.2]) yields Table A5.

Table A5: Partitionings of Total Sums of Squares

(a) Rows before Columns			(b) Columns before Rows		
Term	d.f.	Sum of Squares	Term	d.f.	Sum of Squares
$R(\mu)$	1	392	$R(\mu)$	1	392
$R(\alpha\|\mu)$	1	8	$R(\beta\|\mu)$	2	6
$R(\beta\|\mu,\alpha)$	2	$11\frac{7}{11}$	$R(\alpha\|\mu,\beta)$	1	$13\frac{7}{11}$
$R(\gamma\|\mu,\alpha,\beta)$	2	$36\frac{4}{11}$	$R(\gamma\|\mu,\alpha,\beta)$	2	$36\frac{4}{11}$
SSE	2	10	SSE	2	10
SST	8	458	SST	8	458

Also, the sums of squares from the weighted squares of means analysis are, from (27) and (28)

$$SSA_w = 20 \quad \text{and} \quad SSB_w = 5\tfrac{1}{3}. \tag{A1}$$

The model equations for the data of Table A1 are

$$\begin{bmatrix} 7 \\ 9 \\ 6 \\ 2 \\ 8 \\ 4 \\ 8 \\ 12 \end{bmatrix} = \underset{\sim}{y} = \underset{\sim}{Xb} + \underset{\sim}{e} = \left[\begin{array}{c|cc|ccc|cccccc} 1 & 1 & \cdot & 1 & \cdot & \cdot & 1 & \cdot & \cdot & \cdot & \cdot & \cdot \\ 1 & 1 & \cdot & 1 & \cdot & \cdot & 1 & \cdot & \cdot & \cdot & \cdot & \cdot \\ 1 & 1 & \cdot & \cdot & 1 & \cdot & \cdot & 1 & \cdot & \cdot & \cdot & \cdot \\ 1 & 1 & \cdot & \cdot & \cdot & 1 & \cdot & \cdot & 1 & \cdot & \cdot & \cdot \\ 1 & \cdot & 1 & 1 & \cdot & \cdot & \cdot & \cdot & \cdot & 1 & \cdot & \cdot \\ 1 & \cdot & 1 & \cdot & 1 & \cdot & \cdot & \cdot & \cdot & \cdot & 1 & \cdot \\ 1 & \cdot & 1 & \cdot & 1 & \cdot & \cdot & \cdot & \cdot & \cdot & 1 & \cdot \\ 1 & \cdot & 1 & \cdot & \cdot & 1 & \cdot & \cdot & \cdot & \cdot & \cdot & 1 \end{array}\right] \left[\begin{array}{c} \mu \\ \hline \alpha_1 \\ \alpha_2 \\ \hline \beta_1 \\ \beta_2 \\ \beta_3 \\ \hline \gamma_{11} \\ \gamma_{12} \\ \gamma_{13} \\ \gamma_{21} \\ \gamma_{22} \\ \gamma_{23} \end{array}\right] + \underset{\sim}{e}\,. \qquad \text{(A2)}$$

(Dots in a matrix represent zeros.)

The normal equations (1) resulting from this are

$$\left[\begin{array}{c|cc|ccc|cccccc} 8 & 4 & 4 & 3 & 3 & 2 & 2 & 1 & 1 & 1 & 2 & 1 \\ \hline 4 & 4 & \cdot & 2 & 1 & 1 & 2 & 1 & 1 & \cdot & \cdot & \cdot \\ 4 & \cdot & 4 & 1 & 2 & 1 & \cdot & \cdot & \cdot & 1 & 2 & 1 \\ \hline 3 & 2 & 1 & 3 & \cdot & \cdot & 2 & \cdot & \cdot & 1 & \cdot & \cdot \\ 3 & 1 & 2 & \cdot & 3 & \cdot & \cdot & 1 & \cdot & \cdot & 2 & \cdot \\ 2 & 1 & 1 & \cdot & \cdot & 2 & \cdot & \cdot & 1 & \cdot & \cdot & 1 \\ \hline 2 & 2 & \cdot & 2 & \cdot & \cdot & 2 & \cdot & \cdot & \cdot & \cdot & \cdot \\ 1 & 1 & \cdot & \cdot & 1 & \cdot & \cdot & 1 & \cdot & \cdot & \cdot & \cdot \\ 1 & 1 & \cdot & \cdot & \cdot & 1 & \cdot & \cdot & 1 & \cdot & \cdot & \cdot \\ 1 & \cdot & 1 & 1 & \cdot & \cdot & \cdot & \cdot & \cdot & 1 & \cdot & \cdot \\ 2 & \cdot & 2 & \cdot & 2 & \cdot & \cdot & \cdot & \cdot & \cdot & 2 & \cdot \\ 1 & \cdot & 1 & \cdot & \cdot & 1 & \cdot & \cdot & \cdot & \cdot & \cdot & 1 \end{array}\right] \left[\begin{array}{c} \mu^\circ \\ \hline \alpha_1^\circ \\ \alpha_2^\circ \\ \hline \beta_1^\circ \\ \beta_2^\circ \\ \beta_3^\circ \\ \hline \gamma_{11}^\circ \\ \gamma_{12}^\circ \\ \gamma_{13}^\circ \\ \gamma_{21}^\circ \\ \gamma_{22}^\circ \\ \gamma_{23}^\circ \end{array}\right] = \left[\begin{array}{c} 56 \\ \hline 24 \\ 32 \\ \hline 24 \\ 18 \\ 14 \\ \hline 16 \\ 6 \\ 2 \\ 8 \\ 12 \\ 12 \end{array}\right] \qquad \text{(A3)}$$

with a solution

$$\underset{\sim}{b}^{o\prime} = [0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 8 \quad 6 \quad 2 \quad 8 \quad 6 \quad 12]\,,$$

$$\underset{\sim}{G} = \text{diag}\{0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad \tfrac{1}{2} \quad 1 \quad 1 \quad 1 \quad \tfrac{1}{2} \quad 1\}\,, \qquad \text{(A4)}$$

and

$$\underset{\sim}{b}^{o\prime}\underset{\sim}{X}'\underset{\sim}{y} = [8(16) + 6(6) + 2(2) + 8(8) + 6(12) + 12(12)] = 448.$$

10.2. The generalized "invert part of the inverse" rule

Normal equations for the no-interaction model are, from (A2),

$$\begin{bmatrix} 8 & 4 & 4 & 3 & 3 & 2 \\ 4 & 4 & . & 2 & 1 & 1 \\ 4 & . & 4 & 1 & 2 & 1 \\ 3 & 2 & 1 & 3 & . & . \\ 3 & 1 & 2 & . & 3 & . \\ 2 & 1 & 1 & . & . & 2 \end{bmatrix} \begin{bmatrix} \mu^\circ \\ \alpha_1^\circ \\ \alpha_2^\circ \\ \beta_1^\circ \\ \beta_2^\circ \\ \beta_3^\circ \end{bmatrix} = \begin{bmatrix} 56 \\ 24 \\ 32 \\ 24 \\ 18 \\ 14 \end{bmatrix} \text{ or, equivalently, } \begin{bmatrix} 4 & . & 4 & 2 & 1 & 1 \\ . & 4 & 4 & 1 & 2 & 1 \\ 4 & 4 & 8 & 3 & 3 & 2 \\ 2 & 1 & 3 & 3 & . & . \\ 1 & 2 & 3 & . & 3 & . \\ 1 & 1 & 2 & . & . & 2 \end{bmatrix} \begin{bmatrix} \alpha_1^\circ \\ \alpha_2^\circ \\ \mu^\circ \\ \beta_1^\circ \\ \beta_2^\circ \\ \beta_3^\circ \end{bmatrix} = \begin{bmatrix} 24 \\ 32 \\ 56 \\ 24 \\ 18 \\ 14 \end{bmatrix} . \tag{A5}$$

The second form of the equations in (A5) permits illustration of (17); for example, to calculate $R(\alpha|\mu,\beta)$ using (17), we get $\underset{\sim}{W}$ of (14) as

$$\underset{\sim}{W} = \begin{bmatrix} 4 & . \\ . & 4 \end{bmatrix} - \begin{bmatrix} 4 & 2 & 1 & 1 \\ 4 & 1 & 2 & 1 \end{bmatrix} \begin{bmatrix} . & . & . & . \\ . & \frac{1}{3} & . & . \\ . & . & \frac{1}{3} & . \\ . & . & . & \frac{1}{2} \end{bmatrix} \begin{bmatrix} 4 & 4 \\ 2 & 1 \\ 1 & 2 \\ 1 & 1 \end{bmatrix} = \frac{11}{6} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} . \tag{A6}$$

Then

$$\underset{\sim}{W}^- = \begin{bmatrix} 6/11 & 0 \\ 0 & 0 \end{bmatrix}$$

and it will be found that a solution of the second set of equations in (A5) is

$$\underset{\sim}{b}^\circ = \begin{bmatrix} \alpha_1^\circ \\ \alpha_2^\circ \\ \mu^\circ \\ \beta_1^\circ \\ \beta_2^\circ \\ \beta_3^\circ \end{bmatrix} = \frac{1}{33} \begin{bmatrix} 18 & . & . & -12 & -6 & -9 \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ -12 & . & . & 19 & 4 & 6 \\ -6 & . & . & 4 & 13 & 3 \\ -9 & . & . & 6 & 3 & 21 \end{bmatrix} \begin{bmatrix} 24 \\ 32 \\ 56 \\ 24 \\ 18 \\ 14 \end{bmatrix} = \begin{bmatrix} -2\frac{8}{11} \\ 0 \\ 0 \\ 9\frac{9}{11} \\ 6\frac{10}{11} \\ 8\frac{4}{11} \end{bmatrix} . \tag{A7}$$

We first note that this gives the correct value of the reduction in sum of squares for this model, namely $R(\mu,\alpha,\beta)$: from Table A5

$$R(\mu,\alpha,\beta) = R(\mu) + R(\alpha) + R(\beta|\mu,\alpha) = 392 + 8 + 11\tfrac{7}{11} = 411\tfrac{7}{11} \qquad \text{(A8)}$$

and on applying the R-algorithm of (10) to (A7) we get

$$R(\mu,\alpha,\beta) = [-2\tfrac{8}{11} \quad 0 \quad 0 \quad 9\tfrac{9}{11} \quad 6\tfrac{10}{11} \quad 8\tfrac{4}{11}][24 \quad 32 \quad 56 \quad 24 \quad 18 \quad 14]' = 411\tfrac{7}{11} .$$

Second, applying (17) to (A7) and (A6) gives

$$R(\alpha|\mu,\beta) = [-2\tfrac{8}{11} \quad 0]\begin{bmatrix} 11/6 & -11/6 \\ -11/6 & 11/6 \end{bmatrix}\begin{bmatrix} -2\tfrac{8}{11} \\ 0 \end{bmatrix} = \frac{30^2}{11^2}\left(\frac{11}{6}\right) = \frac{150}{11} = 13\tfrac{7}{11}, \qquad \text{(A9)}$$

the value given in Table A5(b). This is the correct value because (A6) and (A7) have been obtained precisely in accord with (14) and (15) for purposes of deriving $R(\alpha|\mu,\beta)$ from (17). But (17) for $R(\beta|\mu,\alpha)$ does not apply to (A7). If we did try to apply it for $R(\beta|\mu,\alpha)$ we would get

$$[9\tfrac{9}{11} \quad 6\tfrac{10}{11} \quad 8\tfrac{4}{11}]\left[\frac{1}{33}\begin{pmatrix} 19 & 4 & 6 \\ 4 & 13 & 3 \\ 6 & 3 & 21 \end{pmatrix}\right]^{-1}\begin{bmatrix} 9\tfrac{9}{11} \\ 6\tfrac{10}{11} \\ 8\tfrac{4}{11} \end{bmatrix}$$

$$= \frac{16(33)}{121}[27 \;\; 19 \;\; 23]\begin{bmatrix} 19 & 4 & 6 \\ 4 & 13 & 3 \\ 6 & 3 & 21 \end{bmatrix}^{-1}\begin{bmatrix} 27 \\ 19 \\ 23 \end{bmatrix} = \frac{16(33)}{121}[27 \;\; 19 \;\; 23]\frac{1}{132}\begin{bmatrix} 8 & -2 & -2 \\ -2 & 11 & -1 \\ -2 & -1 & 7 \end{bmatrix}\begin{bmatrix} 27 \\ 19 \\ 23 \end{bmatrix}$$

$$= \frac{4}{121}\left\{8(27^2) + 11(19^2) + 7(23^2) - 2[2(27)(19 + 23) + 19(23)]\right\} = \frac{4(8096)}{121} = 267\tfrac{7}{11},$$

which is <u>not</u> $R(\beta|\mu,\alpha)$, the value of which is $11\tfrac{7}{11}$ in Table A5.

10.3. The Σ-restricted model

The Σ-restrictions for the (all-cells-filled) data of Table A1 are as follows:

$$\begin{aligned} \dot{\alpha}_1 + \dot{\alpha}_2 &= 0, & \text{implying } \dot{\alpha}_2 &= -\dot{\alpha}_1 , \\ \dot{\beta}_1 + \dot{\beta}_2 + \dot{\beta}_3 &= 0 & \text{implying } \dot{\beta}_3 &= -\dot{\beta}_1 - \dot{\beta}_2 , \end{aligned} \tag{A10}$$

and

$$\left.\begin{aligned} \dot{\gamma}_{11} + \dot{\gamma}_{12} + \dot{\gamma}_{13} &= 0 \\ \dot{\gamma}_{21} + \dot{\gamma}_{22} + \dot{\gamma}_{23} &= 0 \\ \dot{\gamma}_{11} + \dot{\gamma}_{21} &= 0 \\ \dot{\gamma}_{12} + \dot{\gamma}_{22} &= 0 \\ \dot{\gamma}_{13} + \dot{\gamma}_{23} &= 0 \end{aligned}\right\} \text{ implying } \begin{aligned} \dot{\gamma}_{11} &= \dot{\gamma}_{11} \\ \dot{\gamma}_{12} &= \dot{\gamma}_{12} \\ \dot{\gamma}_{13} &= -\dot{\gamma}_{11} - \dot{\gamma}_{12} \\ \dot{\gamma}_{21} &= -\dot{\gamma}_{11} \\ \dot{\gamma}_{22} &= -\dot{\gamma}_{12} \\ \dot{\gamma}_{23} &= \dot{\gamma}_{11} + \dot{\gamma}_{12} . \end{aligned} \tag{A11}$$

In (A11), the right-hand statements include the obvious $\dot{\gamma}_{11} = \dot{\gamma}_{11}$ and $\dot{\gamma}_{12} = \dot{\gamma}_{12}$. This is to emphasize that the set of restrictions, shown as the left-hand set of statements in (A11), can be restated so that all the $\dot{\gamma}$'s are in terms of just $\dot{\gamma}_{11}$ and $\dot{\gamma}_{12}$. For the general case of a rows and b columns and all cells filled, there will be a + b restrictions on the $\dot{\gamma}$'s, which can be restated so that all $\dot{\gamma}$'s are expressible in terms of just (a-1)(b-1) of them.

The way in which the restrictions change the unrestricted model to the restricted model is seen by applying the restated restrictions of (A10) and (A11) to the model equations (A2). The result is that the model equations for the restricted model are

$$
\begin{bmatrix} 7 \\ 9 \\ 6 \\ 2 \\ 8 \\ 4 \\ 8 \\ 12 \end{bmatrix}
=
\left[\begin{array}{c:c:cc:cc}
1 & 1 & 1 & . & 1 & . \\
1 & 1 & 1 & . & 1 & . \\
1 & 1 & . & 1 & . & 1 \\
1 & 1 & -1 & -1 & -1 & -1 \\
1 & -1 & 1 & . & -1 & . \\
1 & -1 & . & 1 & . & -1 \\
1 & -1 & . & 1 & . & -1 \\
1 & -1 & -1 & -1 & 1 & 1
\end{array}\right]
\begin{bmatrix} \dot{\mu} \\ \dot{\alpha}_1 \\ \dot{\beta}_1 \\ \dot{\beta}_2 \\ \dot{\gamma}_{11} \\ \dot{\gamma}_{12} \end{bmatrix}
+ \underset{\sim}{e} \, .
$$

The normal equations resulting from this are

$$
\left[\begin{array}{c:c:cc:cc}
8 & 0 & 1 & 1 & 1 & -1 \\ \hdashline
0 & 8 & 1 & -1 & 1 & 1 \\ \hdashline
1 & 1 & 5 & 2 & 1 & 0 \\
1 & -1 & 2 & 5 & 0 & -1 \\ \hdashline
1 & 1 & 1 & 0 & 5 & 2 \\
-1 & 1 & 0 & -1 & 2 & 5
\end{array}\right]
\left[\begin{array}{c}
\hat{\dot{\mu}} \\ \hdashline
\hat{\dot{\alpha}}_1 \\ \hdashline
\hat{\dot{\beta}}_1 \\
\hat{\dot{\beta}}_2 \\ \hdashline
\hat{\dot{\gamma}}_{11} \\
\hat{\dot{\gamma}}_{12}
\end{array}\right]
=
\left[\begin{array}{c}
56 \\ \hdashline
-8 \\ \hdashline
10 \\
4 \\ \hdashline
18 \\
4
\end{array}\right] . \qquad \text{(A12)}
$$

a. The no-interaction case

The normal equations for the Σ-restricted, no-interaction case are (A12) with the $\dot{\gamma}$'s and $\dot{\gamma}$-equations removed:

$$
\begin{bmatrix}
8 & 0 & 1 & 1 \\
0 & 8 & 1 & -1 \\
1 & 1 & 5 & 2 \\
1 & -1 & 2 & 5
\end{bmatrix}
\begin{bmatrix} \hat{\dot{\mu}} \\ \hat{\dot{\alpha}}_1 \\ \hat{\dot{\beta}}_1 \\ \hat{\dot{\beta}}_2 \end{bmatrix}
=
\begin{bmatrix} 56 \\ -8 \\ 10 \\ 4 \end{bmatrix}
$$

with solution

$$
\hat{\underset{\sim}{b}} = \frac{1}{22(27)}
\begin{bmatrix}
77 & 0 & -11 & -11 \\
0 & 81 & -27 & 27 \\
-11 & -27 & 152 & -64 \\
-11 & 27 & -64 & 152
\end{bmatrix}
\begin{bmatrix} 56 \\ -8 \\ 10 \\ 4 \end{bmatrix}
=
\begin{bmatrix} 7 \\ -\frac{15}{11} \\ \frac{16}{11} \\ -\frac{16}{11} \end{bmatrix} .
$$

It is easily verified that applying the R-algorithm to this yields $R(\mu,\alpha,\beta)$:

$$R(\mu,\alpha,\beta) = [7 \quad \tfrac{-15}{11} \quad \tfrac{16}{11} \quad \tfrac{-16}{11}][56 \quad -8 \quad 10 \quad 4]' = 411\tfrac{7}{11},$$

as in (A8). Furthermore, by application of (24) it yields $R(\alpha|\mu,\beta)$ and $R(\beta|\mu,\alpha)$:

$$Q_{\dot{\alpha}} = \left(\frac{-15}{11}\right)\left[\frac{81}{22(27)}\right]^{-1}\left(\frac{-15}{11}\right) = \frac{15^2}{11^2}\;\frac{22}{3} = \frac{150}{11} = 13\frac{7}{11} = R(\alpha|\mu,\beta) \text{ of Table A5} \qquad \text{(A13)}$$

and

$$Q_{\dot{\beta}} = \begin{bmatrix}\frac{16}{11} & \frac{-16}{11}\end{bmatrix}\left[\frac{1}{22(27)}\begin{pmatrix}152 & -64\\ -64 & 152\end{pmatrix}\right]^{-1}\begin{bmatrix}16/11\\ 16/11\end{bmatrix} = \left(\frac{16}{11}\right)^2 \frac{22(27)}{8}\,[1 \quad -1]\begin{bmatrix}19 & -8\\ -8 & 19\end{bmatrix}^{-1}\begin{bmatrix}1\\ -1\end{bmatrix}$$

$$= \frac{64}{11^2}\,(19 + 19 - 16) = \frac{128}{11} = 11\frac{7}{11} = R(\beta|\mu,\alpha) \text{ of Table A5.}$$

b. The interaction case

The solution to equation (A12) is

$$\hat{\underset{\sim}{b}} = \frac{1}{72}\begin{bmatrix}10 & 0 & -1 & -1 & -3 & 3\\ 0 & 10 & -3 & 3 & -1 & -1\\ -1 & -3 & 19 & -8 & -3 & 0\\ -1 & 3 & -8 & 19 & 0 & 3\\ -3 & -1 & -3 & 0 & 19 & -8\\ 3 & -1 & 0 & 3 & -8 & 19\end{bmatrix}\begin{bmatrix}56\\ -8\\ 10\\ 4\\ 18\\ 4\end{bmatrix} = \begin{bmatrix}7\\ -10/6\\ 1\\ -1\\ 10/6\\ 10/6\end{bmatrix}, \qquad \text{(A14)}$$

from which we may verify $R(\mu,\alpha,\beta,\gamma)$: from Table A5

$$R(\mu,\alpha,\beta,\gamma) = SST - SSE = 458 - 10 = 448$$

and from applying the R-algorithm to (A14)

$$R(\dot{\mu},\dot{\alpha},\dot{\beta},\dot{\gamma}) = R(\mu,\alpha,\beta,\gamma)$$

$$= [7 \quad \tfrac{-10}{6} \quad 1 \quad -1 \quad \tfrac{10}{6} \quad \tfrac{-10}{6}][56 \quad -8 \quad 10 \quad 4 \quad 18 \quad 4]' = 448.$$

We can also illustrate (31):

$$Q_{\dot{\alpha}} = \left(\frac{-10}{6}\right)\left(\frac{10}{72}\right)^{-1}\left(\frac{-10}{6}\right) = \frac{120}{6} = 20 = SSA_w \text{ of (A1)},$$

and

$$Q_{\dot{\beta}} = [1 \quad -1]\left[\frac{1}{72}\begin{pmatrix}19 & -8\\ -8 & 19\end{pmatrix}\right]^{-1}\begin{bmatrix}1\\ -1\end{bmatrix} = \frac{72}{11(27)}\ (19 + 19 - 16) = 5\tfrac{1}{3} = SSB_w \text{ of (A1)}.$$

Finally, from (A14) we can also verify $R(\gamma|\mu,\alpha,\beta)$ using (31):

$$Q_{\dot{\gamma}} = \begin{bmatrix}\frac{10}{6} & \frac{10}{6}\end{bmatrix}\left[\frac{1}{72}\begin{pmatrix}19 & -8\\ -8 & 19\end{pmatrix}\right]^{-1}\begin{bmatrix}10/6\\ 10/6\end{bmatrix} = \frac{100}{36}\,\frac{72}{11(27)}\ (19 + 19 + 16) = \frac{400}{11} = 36\,\frac{4}{11}$$

$$= R(\gamma|\mu,\alpha,\beta) \text{ of Table A5.}$$

We can also illustrate calculation of $R^*(\alpha|\mu,\beta,\gamma)_\Sigma$ of (59) and (60). It will, of course, equal SSA_w in this case — because all cells are filled. For (A12), equations involving $\dot{\mu}$, $\dot{\beta}$ and $\dot{\gamma}$ are

$$\begin{bmatrix}8 & 1 & 1 & 1 & -1\\ 1 & 5 & 2 & 1 & 0\\ 1 & 2 & 5 & 0 & -1\\ 1 & 1 & 0 & 5 & 2\\ -1 & 0 & -1 & 2 & 5\end{bmatrix}\begin{bmatrix}\hat{\dot{\mu}}\\ \hat{\dot{\beta}}_1\\ \hat{\dot{\beta}}_2\\ \hat{\dot{\gamma}}_{11}\\ \hat{\dot{\gamma}}_{12}\end{bmatrix} = \begin{bmatrix}56\\ 10\\ 4\\ 18\\ 4\end{bmatrix} \text{ with solution } \begin{bmatrix}7\\ \frac{1}{2}\\ -\frac{1}{2}\\ 1\frac{1}{2}\\ 1\frac{1}{2}\end{bmatrix}. \qquad \text{(A15)}$$

Therefore on applying the R-algorithm to (A15)

$$R^*(\dot{\mu},\dot{\beta},\dot{\gamma})_\Sigma = [7 \quad \tfrac{1}{2} \quad -\tfrac{1}{2} \quad 1\tfrac{1}{2} \quad 1\tfrac{1}{2}][56 \quad 10 \quad 4 \quad 18 \quad 4]' = 428$$

and so, as in (59) and (60),

$$R^*(\dot{\alpha}|\dot{\mu},\dot{\beta},\dot{\gamma})_\Sigma = R(\mu,\alpha,\beta,\gamma) - R^*(\dot{\mu},\dot{\beta},\dot{\gamma})_\Sigma = 448 - 428 = 20 = SSA_w\ .$$

Consider $R(\mu,\beta,\gamma)$ for the example: it comes from the model $E(y_{ijk}) = \mu + \beta_j + \gamma_{ij}$; and since there are 6 cells with data, the rank of the normal equations for that model is 6. Now notice that equations (A15), which yield $R^*(\dot{\mu},\dot{\beta},\dot{\gamma})_\Sigma$, have rank 5. This illustrates why $R^*(\dot{\mu},\dot{\beta},\dot{\gamma})_\Sigma \neq R(\mu,\beta,\gamma)$, as stated in (56).

10.4. The 0-restrictions

Suppose the 0_{11}-restrictions corresponding to (62) are:

$$\begin{array}{llll} \dot{\alpha}_1 = 0 & \dot{\gamma}_{11} = 0 & \dot{\gamma}_{12} = 0 & \dot{\gamma}_{13} = 0 \\ \dot{\beta}_1 = 0 & \dot{\gamma}_{21} = 0 . & & \end{array} \tag{A16}$$

The effect of these on the model equations (A2) of the unrestricted model is to eliminate columns of the $\underset{\sim}{X}$-matrix corresponding to the elements equated to zero in (A16). The effect on the normal equations (A3) is deletion of the corresponding rows and columns, thus giving the normal equations for the restricted model as

$$\left[\begin{array}{c|c|cc|cc} 8 & 4 & 3 & 2 & 2 & 1 \\ \hline 4 & 4 & 2 & 1 & 2 & 1 \\ \hline 3 & 2 & 3 & 0 & 2 & 0 \\ 2 & 1 & 0 & 2 & 0 & 1 \\ \hline 2 & 2 & 2 & 0 & 2 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 \end{array}\right] \left[\begin{array}{c} \hat{\dot{\mu}} \\ \hline \hat{\dot{\alpha}}_2 \\ \hline \hat{\dot{\beta}}_2 \\ \hat{\dot{\beta}}_3 \\ \hline \hat{\dot{\gamma}}_{22} \\ \hat{\dot{\gamma}}_{23} \end{array}\right] = \left[\begin{array}{c} 56 \\ \hline 32 \\ \hline 18 \\ 14 \\ \hline 12 \\ 12 \end{array}\right] . \tag{A17}$$

The solution to this is

$$\begin{bmatrix} \hat{\dot{\mu}} \\ \hat{\dot{\alpha}}_2 \\ \hat{\dot{\beta}}_2 \\ \hat{\dot{\beta}}_3 \\ \hat{\dot{\gamma}}_{22} \\ \hat{\dot{\gamma}}_{33} \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 & -1 & -1 & -1 & 1 & 1 \\ -1 & 3 & 1 & 1 & -3 & -3 \\ -1 & 1 & 3 & 1 & -3 & -1 \\ -1 & 1 & 1 & 3 & -1 & -3 \\ 1 & -3 & -3 & -1 & 6 & 3 \\ 1 & -3 & -1 & -3 & 3 & 7 \end{bmatrix} \begin{bmatrix} 56 \\ 32 \\ 18 \\ 14 \\ 12 \\ 12 \end{bmatrix} = \begin{bmatrix} 8 \\ 0 \\ -2 \\ -6 \\ 0 \\ 10 \end{bmatrix} .$$

Using the R-algorithm on this, the value of $R(\mu,\alpha,\beta,\gamma)$ is again readily verified:

$$R(\mu,\alpha,\beta,\gamma) = R(\dot{\mu},\dot{\alpha},\dot{\beta},\dot{\gamma})_0 = [8 \quad 0 \quad -2 \quad -6 \quad 0 \quad 10][56 \quad 32 \quad 18 \quad 14 \quad 12 \quad 12]' = 448 \ .$$

Then

$$R^*(\dot{\alpha}|\dot{\mu},\dot{\beta},\dot{\gamma})_{0_{11}} = Q_{\dot{\alpha}} = 0(3/2)^{-1}0 = 0$$

and

$$R^*(\dot{\beta}|\dot{\mu},\dot{\alpha},\dot{\gamma})_{0_{11}} = Q_{\dot{\beta}} = (-2 \quad -6)\left[\frac{1}{2}\begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix}\right]^{-1}\begin{pmatrix} -2 \\ -6 \end{pmatrix} \tag{A18}$$

$$= 8(1 \quad 3)\,\frac{1}{8}\begin{pmatrix} 3 & -1 \\ -1 & 3 \end{pmatrix}\begin{pmatrix} 1 \\ 3 \end{pmatrix} = 24 \ . \tag{A19}$$

Confirmation of this latter result comes from noting from (61) that

$$\text{using } R^*(\dot{\beta}|\dot{\mu},\dot{\alpha},\dot{\gamma})_{0_{11}} \text{ tests } H: \beta_1 + \gamma_{11} = \beta_2 + \gamma_{12} = \beta_3 + \gamma_{13} \ .$$

Writing this hypothesis (in the unrestricted model) is

$$H: \underset{\sim}{K}'\underset{\sim}{b} = 0 \quad \text{for} \quad \underset{\sim}{K}' = \begin{bmatrix} 0 & 0 & 0 & 1 & -1 & 0 & 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & -1 & 1 & 0 & -1 & 0 & 0 & 0 \end{bmatrix}. \tag{A20}$$

Then from (A4)

$$\underset{\sim}{K}'\underset{\sim}{b}^0 = \begin{bmatrix} 8-6 \\ 8-2 \end{bmatrix} = \begin{bmatrix} 2 \\ 6 \end{bmatrix} \quad \text{and} \quad \underset{\sim}{K}'\underset{\sim}{G}\underset{\sim}{K} = \begin{bmatrix} \frac{1}{2} & -1 & 0 \\ \frac{1}{2} & 0 & -1 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ -1 & 0 \\ 0 & -1 \end{bmatrix} = \begin{bmatrix} \frac{3}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{3}{2} \end{bmatrix}.$$

Hence the numerator sum of squares (12) for testing the hypothesis of (A20) is

$$Q = [2 \quad 6]\begin{bmatrix} \frac{3}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{3}{2} \end{bmatrix}^{-1}\begin{bmatrix} 2 \\ 6 \end{bmatrix}$$

which is precisely (A18), i.e., $R^{*}(\dot{\beta}|\dot{\mu},\dot{\alpha},\dot{\gamma})_{0_{11}} = 24$.

10.5. The W-restrictions

The W-restrictions for the illustrative data are as follows:

$$4\dot{\alpha}_1 + 4\dot{\alpha}_2 = 0, \quad \text{implying} \quad \dot{\alpha}_2 = -\dot{\alpha}_1 \tag{A21}$$

$$3\dot{\beta}_1 + 3\dot{\beta}_2 + 2\dot{\beta}_3 = 0, \quad \text{implying} \quad \dot{\beta}_3 = -1\tfrac{1}{2}(\dot{\beta}_1 + \dot{\beta}_2) \tag{A22}$$

and

$$\begin{aligned}
2(\dot{\alpha}_1 + \dot{\gamma}_{11}) + \dot{\alpha}_2 + \dot{\gamma}_{21} &= 0 \\
\dot{\alpha}_1 + \dot{\gamma}_{12} + 2(\dot{\alpha}_2 + \dot{\gamma}_{22}) &= 0 \\
\dot{\alpha}_1 + \dot{\gamma}_{13} + \dot{\alpha}_2 + \dot{\gamma}_{23} &= 0 \\
2(\dot{\beta}_1 + \dot{\gamma}_{11}) + \dot{\beta}_2 + \dot{\gamma}_{12} + \dot{\beta}_3 + \dot{\gamma}_{13} &= 0 \\
\dot{\beta}_1 + \dot{\gamma}_{21} + 2(\dot{\beta}_2 + \dot{\gamma}_{22}) + \dot{\beta}_3 + \dot{\gamma}_{23} &= 0
\end{aligned}$$

all of which imply, in the nature of (A11),

$$\begin{aligned}
\dot{\gamma}_{11} &= \dot{\gamma}_{11} \\
\dot{\gamma}_{12} &= \dot{\gamma}_{12} \\
\dot{\gamma}_{13} &= -(\tfrac{1}{2}\dot{\beta}_1 - \tfrac{1}{2}\dot{\beta}_2 + 2\dot{\gamma}_{11} + \dot{\gamma}_{12} \\
\dot{\gamma}_{21} &= -(\dot{\alpha}_1 + 2\dot{\gamma}_{11}) \\
\dot{\gamma}_{22} &= \tfrac{1}{2}(\dot{\alpha}_1 - \dot{\gamma}_{12})
\end{aligned} \tag{A23}$$

and

$$\dot{\gamma}_{23} = \tfrac{1}{2}\dot{\beta}_1 - \tfrac{1}{2}\dot{\beta}_2 + 2\dot{\gamma}_{11} + \dot{\gamma}_{12} .$$

Applying (A21), (A22) and (A23) to the model equations (A2) of the unrestricted model yields the model equations for this restricted model, with W-restrictions, as

$$\left[\begin{array}{c|c|cc|cc} 1 & 1 & 1 & \cdot & 1 & \cdot \\ 1 & 1 & 1 & \cdot & 1 & \cdot \\ 1 & 1 & \cdot & 1 & \cdot & 1 \\ 1 & 1 & -2 & -1 & -2 & -1 \\ 1 & -2 & 1 & \cdot & -2 & \cdot \\ 1 & -\tfrac{1}{2} & \cdot & 1 & \cdot & -\tfrac{1}{2} \\ 1 & -\tfrac{1}{2} & \cdot & 1 & \cdot & -\tfrac{1}{2} \\ 1 & -1 & -1 & -2 & 2 & 1 \end{array}\right] \begin{bmatrix} \dot{\mu} \\ \dot{\alpha}_1 \\ \dot{\beta}_1 \\ \dot{\beta}_2 \\ \dot{\gamma}_{11} \\ \dot{\gamma}_{12} \end{bmatrix} = \begin{bmatrix} 7 \\ 9 \\ 6 \\ 2 \\ 8 \\ 4 \\ 8 \\ 12 \end{bmatrix} . \tag{A24}$$

The corresponding normal equations are then

$$\left[\begin{array}{c|c|cc|cc} 8 & \cdot & \cdot & \cdot & \cdot & \cdot \\ \hline \cdot & 9\tfrac{1}{2} & -1 & 1 & 2 & -\tfrac{1}{2} \\ \hline \cdot & -1 & 8 & 4 & 2 & 1 \\ \cdot & 1 & 4 & 8 & -2 & -1 \\ \hline \cdot & 2 & 2 & -2 & 14 & 4 \\ \cdot & -\tfrac{1}{2} & 1 & -1 & 4 & 3\tfrac{1}{2} \end{array}\right] \left[\begin{array}{c} \hat{\dot{\mu}} \\ \hline \hat{\dot{\alpha}}_1 \\ \hline \hat{\dot{\beta}}_1 \\ \hat{\dot{\beta}}_2 \\ \hline \hat{\dot{\gamma}}_{11} \\ \hat{\dot{\gamma}}_{12} \end{array}\right] = \left[\begin{array}{c} 56 \\ \hline -10 \\ \hline 8 \\ -8 \\ \hline 20 \\ 10 \end{array}\right] . \tag{A25}$$

Instead of using the $Q_{\dot{\alpha}}$ algorithm of (22) we simply solve (A25) to get

$$\hat{\dot{\underset{\sim}{b}}}' = [7 \quad -1 \quad 1 \quad -1 \quad 1 \quad 1];$$

and then the R-algorithm gives

$$R(\dot{\mu},\dot{\alpha},\dot{\beta},\dot{\gamma})_W = 7(56) + 10 + 8 + 8 + 20 + 10 = 448 = R(\mu,\alpha,\beta,\gamma),$$

as expected.

To demonstrate (65), delete β's and β-equations from (A25) leaving

$$\begin{bmatrix} 8 & \cdot & \cdot & \cdot \\ \cdot & 9\frac{1}{2} & 2 & -\frac{1}{2} \\ \cdot & 2 & 14 & 4 \\ \cdot & -\frac{1}{2} & 4 & 3\frac{1}{2} \end{bmatrix} \begin{bmatrix} \hat{\dot{\mu}} \\ \hat{\dot{\alpha}}_1 \\ \hat{\dot{\gamma}}_{11} \\ \hat{\dot{\gamma}}_{12} \end{bmatrix} = \begin{bmatrix} 56 \\ -10 \\ 20 \\ 10 \end{bmatrix} \quad \text{with solution} \quad \begin{bmatrix} 7 \\ -1\frac{1}{4} \\ 1\frac{1}{4} \\ 1\frac{1}{4} \end{bmatrix}.$$

Then the R-algorithm gives

$$R^*(\dot{\mu},\dot{\alpha},\dot{\gamma})_W = 7(56) - 1\tfrac{1}{4}(-10) + 1\tfrac{1}{4}(20) + 1\tfrac{1}{4}(10) = 442\,.$$

Hence, similar to (59),

$$R^*(\dot{\beta}|\dot{\mu},\dot{\alpha},\dot{\gamma})_W = R(\mu,\alpha,\beta,\gamma) - R^*(\dot{\mu},\dot{\alpha},\dot{\gamma})_W = 448 - 442 = 6 = R(\beta|\mu) \text{ of Table A5,}$$

so illustrating (65). The reader can use (A25) to verify $R(\alpha|\mu)$ in the same manner.

11. Appendix B: Deriving SSA_w from Testing a Hypothesis

To use

$$Q = \underset{\sim}{b}^{o\prime}\underset{\sim}{K}(\underset{\sim}{K}'\underset{\sim}{G}\underset{\sim}{K})^{-1}\underset{\sim}{K}'\underset{\sim}{b}^{o} \qquad \text{(B1)}$$

of (12) for the hypothesis of (26), namely $H: \alpha_i + \bar{\gamma}_{i.}$ all equal, we first rewrite it in the form

$$H: \mu + \alpha_1 + \bar{\beta}_. + \bar{\gamma}_{1.} = \mu + \alpha_i + \bar{\beta}_. + \bar{\gamma}_{i.} \quad \text{for } i = 2, \cdots, a. \qquad \text{(B2)}$$

This is a testable hypothesis because $\mu + \alpha_i + \bar{\beta}_. + \bar{\gamma}_{i.} = \sum_{j=1}^{b} (\mu + \alpha_i + \beta_j + \gamma_{ij})/a$, in which the expression in parentheses is estimable — and $\bar{\beta}_.$ occurs in (B2) for all i because of having all cells filled. Now recall that for the model (25) the only non-null parts of $\underset{\sim}{b}^o$ and $\underset{\sim}{G}$ are the vector $\{\bar{y}_{ij.}\}$ of cell means and the diagonal matrix $\underset{\sim}{D}\{1/n_{ij}\}$, respectively, corresponding to the elements γ_{ij} of $\underset{\sim}{b}$. Hence for the hypothesis (B2) the terms of Q in (B1) are

$$\underset{\sim}{K}'\underset{\sim}{b}^{o} = \underset{\sim}{L}'\{\bar{y}_{ij.}\} \quad \text{and} \quad \underset{\sim}{K}'\underset{\sim}{G}\underset{\sim}{K} = \underset{\sim}{L}'\underset{\sim}{D}\{1/n_{ij}\}\underset{\sim}{L}$$

where

$$\underset{\sim}{L}' = \frac{1}{b}\left[\underset{\sim}{1}_{a-1} \,\vdots\, -\underset{\sim}{I}_{a-1}\right] \otimes \underset{\sim}{1}_b' \qquad \text{(B3)}$$

with $\underset{\sim}{1}_b'$ being a row vector of b unities and $\otimes$ being the Kronecker product operator. Hence

$$\underset{\sim}{K}'\underset{\sim}{b}^{o} = \frac{1}{b}\left\{\sum_{j=1}^{b} \bar{y}_{1j.} - \sum_{j=1}^{b} \bar{y}_{ij.}\right\} = \left\{\bar{x}_{1.} - \bar{x}_{i.}\right\} \quad \text{for } i = 2, \cdots, a$$

on using (28); and

$\underset{\sim}{K}'\underset{\sim\sim}{GK}$ has diagonal elements $\frac{1}{b^2}\sum_{j=1}^{b}\left(\frac{1}{n_{1j}}+\frac{1}{n_{ij}}\right) = \frac{1}{w_1}+\frac{1}{w_i}$, for $i = 2, \cdots, a$

and off-diagonal elements $\frac{1}{b^2}\sum_{j=1}^{b}\frac{1}{n_{1j}} = \frac{1}{w_1}$.

Therefore

$$(\underset{\sim}{K}'\underset{\sim\sim}{GK})^{-1} = \left[\underset{\sim}{D}\{\frac{1}{w_i}\} + \frac{1}{w_1}\underset{\sim\sim}{\mathbf{11}}'\right]^{-1} \text{ for } i = 2, \cdots, a, \text{ and } \underset{\sim}{\mathbf{1}}' \text{ of order } a-1$$

$$= \underset{\sim}{D}^{-1}\{\frac{1}{w_i}\} - \frac{\frac{1}{w_1}}{1+\frac{1}{w_1}\underset{\sim}{\mathbf{1}}'\underset{\sim}{D}^{-1}\{\frac{1}{w_i}\}\underset{\sim}{\mathbf{1}}}\underset{\sim}{D}^{-1}\{\frac{1}{w_i}\}\underset{\sim\sim}{\mathbf{11}}'\underset{\sim}{D}^{-1}\{\frac{1}{w_i}\}$$

$$= \underset{\sim}{D}\{w_i\} - \frac{1}{w_.}\{w_iw_{i'}\} \text{ for } i,i' = 2, \cdots, a \text{ and } w_. = \sum_{i=1}^{a} w_i .$$

Hence

$$Q = \{\bar{x}_{1.} - \bar{x}_{i.}\}'[\underset{\sim}{D}\{w_i\} - \frac{1}{w_.}\{w_iw_{i'}\}]\{\bar{x}_{1.} - \bar{x}_{i.}\} \text{ for } i,i' = 2, \cdots, a$$

and, after some tedious algebra, this simplifies to

$$Q = \sum_{i=1}^{a} w_i(\bar{x}_{i.} - \bar{x}_{[1]})^2 = SSA_w .$$